

Future shock

At first glance, India's electorate has just put the brakes on the country's modernization, by ousting a government that has promoted entrepreneurship and invested in high technology. But appearances can be deceptive.

Few observers anticipated last week's impressive Indian election victory by the Congress Party, led by Italian-born Sonia Gandhi, widow of slain prime minister Rajiv Gandhi.

The defeated government of the nationalist Bharatiya Janata Party (BJP) was credited with nurturing rapid economic growth in India and attracting foreign investment. During its rule, according to some estimates, 250,000 expatriates returned home, many of them scientists and engineers drawn into the booming biotechnology and information-technology sectors. In the past couple of years, the economy has expanded at an unprecedented annual rate of around 8%.

The BJP government increased the amount that India spends on research and development from 0.6% of gross domestic product in 1999 to 1.1% today — and it claimed that this would rise to 2% by 2007. It also supported ambitious industrial ventures, including the beginnings of a civilian aircraft industry, the arrival of telemedicine in remote villages, and the expansion of higher education.

But many of the reforms that laid the groundwork for these successes were initiated by the Congress Party before it lost power in 1998. Despite the panic in the financial markets as Gandhi negotiated with the left-wing parties whose support will be needed to form a stable government, there is little reason to fear that the economic expansion will end with the rule of the BJP.

Support for scientific research is likely to be largely maintained by the incoming government. The Congress Party's traditional rationalism will also be welcomed by many scientists, who may be glad to see the back of former BJP science minister Murli Manohar Joshi, who sought to persuade universities to teach astrology (see *Nature* **411**, 227; 2001). Joshi was defeated at the polls in Uttar Pradesh.

Nevertheless, the new government must address a series of important science-related challenges. The BJP embraced technology as the

solution to many national problems. Its manifesto promised to usher in a second 'green revolution' with the help of biotechnology, and to engage in substantial civil-engineering projects that would link up rivers to improve water supplies (see *Nature* **422**, 790; 2003).

Both approaches were criticized by environmentalists, who will hold more sway with the new government. A more cautious approach can therefore be expected in pursuing agricultural biotechnology in particular. Gandhi must ensure that this caution doesn't leave India's farmers grappling with outdated seeds, tools and methods.

One of India's next major appointments on the international stage will arise next spring, when the nuclear non-proliferation treaty is due for renewal. The Congress Party has said that it will maintain the nuclear weapons that the BJP tested in 1998, but can be expected to revive India's role in pressing states with nuclear weapons to fulfil their commitments under the treaty.

India also holds a pivotal position on the World Trade Organization's Trade-Related Aspects of Intellectual Property Rights (TRIPS) agreement, which is supposed to harmonize patent rules. India is due to comply with TRIPS by January next year. But it will be interesting to see how the necessary legislation fares in the new parliament, given that Congress's supporters will be hurt by reforms that could undermine India's generic-drugs industry and increase the cost of healthcare.

This conflict between market-friendly reform and the grim reality of life in India's villages and cities partly explains last week's electoral shock. It is a good thing that the conflict is being resolved through the ballot box and not the barrel of a gun. Investors are not always rational — but if they are, the message they will take home from last week's smooth transfer of power is that India's young institutions are well equipped to underpin the country's development. ■

Equal treatment under the law

If this fundamental right is denied to visiting researchers, the quality of US science will suffer.

Since February 2003, Iranian physicist Shahram Rahatlou has been barred from entering the government lab where he worked. The security officials who came to his door never told him what he had done wrong. They said there would be no appeal.

The lab is not in Iran, but at the Stanford Linear Accelerator Center (SLAC) in California. Rahatlou had been working there for nearly five years when officers from the US Department of Energy (DOE) visited him. He had, by all accounts, been an exceptional physicist. SLAC is a purely academic laboratory, which houses no classified research, and the experiment on which he was working makes all its findings publicly available. Even the lab's scientific leaders don't know why Rahatlou was barred from the premises (see page 229).

The DOE, which manages nuclear weapons facilities in addition to running civilian laboratories such as SLAC, has a duty to protect national security. Since the attacks of 11 September 2001, it has subjected many foreign researchers to background checks. But in a democracy, those who come under suspicion should have the right

to know the accusations that they face, and have the right to appeal.

Rahatlou may come from a state designated as a "sponsor of terror", but he is also one of the most hard-working physicists in a collaboration of more than 600 researchers from 10 countries. It is time for the US government to take a second look at his case. The DOE says the matter cannot be discussed further because it is classified, but many scientific officials — including John Marburger, the president's science adviser — have high-level security clearance. It is up to them to find a way to give Rahatlou his day in court.

The shocking pictures that have emerged from the Abu Ghraib jail in Iraq have sullied the United States' reputation across the Islamic world. Rahatlou's case is not in the same league of infamy. But it adds to a damaging impression that US officials are prepared to ride roughshod over those they see as potential enemies in the 'war on terror'. If US labs are to continue to attract the world's best scientific talent, visitors need to know that they will be treated with respect and dignity under the law. ■





On the attack
Congress plans
deeper probe of
NIH conflicts
p230



Computer wars
Software ban sparks
heated reaction
from chemists
p231



Wake-up call
America braces
itself for a cicada
invasion
p233



Hard to swallow
Liver lovers bid to
put the puffer fish
back on the menu
p234

Iranian physicist locked out of laboratory by energy department

Geoff Brumfiel, Washington

His supervisor believes he will be one of the achievers of tomorrow, he has just won an award from the American Physical Society — and he has been banned from his lab for more than a year by the US Department of Energy (DOE).

Shahram Rahatlou is an Iranian high-energy physicist based at the University of California, San Diego (UCSD), whose research involves using the BaBar particle detector at the Stanford Linear Accelerator Center (SLAC). But since February last year, he has been prohibited from entering the grounds of SLAC, where he has worked for the past six years.

Rahatlou declined to comment for this article, but colleagues and laboratory sources told *Nature* that security officials from the DOE, which oversees SLAC, have offered no explanation for his expulsion and have stated that there will be no appeal.

SLAC is considered to be a civilian facility and does not conduct classified or sensitive research. But security rules have been tightened at such facilities, in part because the DOE is also responsible for research into sensitive topics such as nuclear weapons.

Vivek Sharma, a UCSD physicist and Rahatlou's thesis adviser, says that the 30-year-old has not violated any immigration laws, and that his work on BaBar has no bearing on either national security or weapons of mass destruction.

Rahatlou, who is working in the United States on the type of visa — an H-1B — used by many visiting scientists, has continued to work remotely on the BaBar experiment from his office at UCSD. But the ruling effectively puts his career on hold, says Sharma. "If the current isolation remains, then I think he has to leave the field," he says.

Rahatlou's family moved from Iran to Rome in 1987, where he attended the city's 'La Sapienza' university. It was there that he caught the eye of physicist Fernando Ferroni, who believed that Rahatlou showed great promise. "He was very bright," Ferroni recalls. On Ferroni's recommendation, Rahatlou joined Sharma's group in the autumn of 1998. "He's probably the best graduate stu-



No entry: Shahram Rahatlou (below) has been unable to work at this Stanford lab for more than a year.

dent I have had," Sharma says.

In four years, much of it spent in an office at SLAC, Rahatlou completed his PhD thesis on how measurements made with BaBar might help to explain why there is vastly more matter than antimatter in the Universe. His work won him the 2004 Mitsuyoshi Tanaka Dissertation Award, an annual prize for an exceptional thesis in experimental particle physics. The award was presented at this month's American Physical Society meeting in Denver.

His troubles began after he completed his PhD in the autumn of 2002. Rahatlou secured a postdoctoral fellowship at Fermilab near Chicago, Illinois. But to take the position, he had to undergo a background check as part of security procedures introduced after the terrorist attacks of 11 September 2001. Sharma suspects that it was this that led two DOE officials to his UCSD lab in February last year. Sharma says that the officials told him that Rahatlou was no longer welcome at DOE facilities, but that they did not explain why. They also said that the case could not be appealed. Sharma's version of events has been confirmed by several sources in the physics community, who declined to be identified.

A laboratory spokesman says that offi-



cials at SLAC, which is run by Stanford University for the DOE, have also not received an explanation for the ruling. "It was a DOE decision," he says.

Other Iranians say that the new security clearances have made it harder for them to work at, or even visit, government facilities. "We were discriminated against even before 11 September," says Niayesh Afshordi, an Iranian graduate student at Princeton University in New Jersey. "But now it has become worse." Afshordi, who expects to complete his PhD in theoretical astrophysics later this year, says that he had applied for a postdoctoral position at Fermilab, but a fellow researcher there told him that the on-site security clearance would take too long to make it worthwhile.

DOE officials who oversee SLAC failed to return calls enquiring about the decision. Jeanne Lopatto, a spokeswoman at the department's Washington headquarters, said in an e-mail: "DOE will decline to comment on this."

Sharma says he hopes that the DOE will review the ruling. "Rahatlou is the kind of guy who was going on to be the next brilliant faculty at one of the top US universities," he says. "The chances of that happening are dimming day-by-day."

SLAC

APS

Health assembly rebuffs Taiwan's bid for 'observer' status

Declan Butler

As officials from 192 countries gathered in Geneva this week for the annual meeting of the World Health Assembly, Chien-Jen Chen, Taiwan's minister of health, was left knocking on the door.

Taiwan is not recognized by the World Health Organization (WHO) — its access to WHO people and programmes is at the discretion of China. But Chen's government wants to be given 'observer' status, arguing that its scientists are currently being shut out of global investigations into severe acute respiratory syndrome (SARS) and other diseases (see *Nature* 422, 652; 2003).

In the event, Taiwan's bid did not get very far. The assembly's general committee decided not to put it on the agenda. The country's eighth unsuccessful bid was rejected after what one WHO official describes as "lengthy discussions", adding that Taiwan did secure strong US and Japanese backing for the first time.

China does not recognize the government of Taiwan. The United Nations, which admitted China in 1972, also does not consider Taiwan a state, which prevents Taiwan from becoming a full member of WHO. But the World Health Assembly, which serves as WHO's main decision-making body, could grant Taiwan observer status, if it so wished.

Chen says that the bid "tried to avoid one-China politics", and describes it as Taiwan's "humble request to help realize health for all". But in a report by China's official Xinhua news agency on Monday, Gao Qiang, the Chinese vice-minister of health, claims Taiwan "has no difficulties in getting information from WHO and carrying out technical exchanges with it". Taiwan's true purpose, Qiang is quoted as saying, is to "politicize the health issue and to internationalize the Taiwan issue".

Qiang adds that China has permitted Taiwan experts to attend WHO meetings. But Chen counters that they were invited to SARS meetings on the basis of their expertise and that Chinese interference delayed the invitations.

According to a WHO official, many of its members, including most European countries, do not want to grant Taiwan observer status in case this is seen as a step towards diplomatic recognition. But Chen says he will be back next year to press the case for the admission of a country that, with 23 million people, has a larger population than three-quarters of the WHO member states. ■

Panel slated for leniency over study of NIH consulting roles

Erika Check, Washington

A special panel that investigated conflict-of-interest practices at the National Institutes of Health (NIH) got a chilly reception on Capitol Hill last week. Law-makers and their staffs made it clear that the biomedical agency will be dogged by the issue for many months to come.

The 'blue ribbon panel' was chaired by Bruce Alberts, president of the National Academy of Sciences, and Norman Augustine, chairman of the defence contractor Lockheed Martin. The pair joined NIH director Elias Zerhouni to brief law-makers at a hearing on 12 May. This was just six days after the panel released its report recommending changes to the way that the NIH manages the consulting partnerships its employees have with private companies and academic institutions (see *Nature* 429, 119; 2004).

But members of Congress at the hearing said the panel had not gone nearly far enough, and that they would continue to probe the NIH's practices. "It is clear that some NIH scientists are very close to the line or have crossed the line," said James Greenwood (Republican, Pennsylvania), chairman of the investigations subcommittee of the House of Representatives' Committee on Energy and Commerce. "The blue ribbon panel seems to handle conflict-of-interest issues gently."

And the chairman of the full energy and commerce committee, Joe Barton (Republican, Texas), accused the NIH and its parent agency, the Department of Health and Human Services, of stalling Greenwood's investigation. "We have found the NIH to be less than cooperative," Barton said. "Your officials can be cooperative cooperatively, or we're going to make them cooperative coercively," Barton told Zerhouni.

One veteran congressional staff member says the conflict-of-interest issue has put "a huge black mark" against the NIH. The agency has recently been favoured by law-makers of both parties, who doubled its budget from 1998 to 2003. Now, the staff member says, "the bloom is off the rose at the NIH".

At the hearing, law-makers appeared incensed that NIH scientists had been engaging in practices that they considered to be

ethically indefensible while the agency was receiving huge budget increases. According to congressional staff and lobbyists, some on Capitol Hill feel personally betrayed by the revelations about the NIH's outside consulting deals, which first appeared in the *Los Angeles Times* last December.

And although congressional committees steered clear of overseeing the NIH during the period of its budget doubling, their approach is now changing dramatically. At the 12 May hearing, Barton indicated that he would try this year to pass a law called a



James Greenwood: says some NIH scientists have "crossed the line".

"reauthorization", which will spell out policies and procedures for the NIH. Such bills are supposed to be passed every five years or so, but Congress has not reauthorized the agency since 1993.

"You have got members of Congress who were not really engaged in oversight who now want to step up to the plate," says one NIH advocate. Greenwood's subcommittee was due to hold a second hearing on the issue on 18 May, six days after the first one.

Biomedical lobbyists predict that the conflict-of-interest issue will follow the NIH for a while. One said that it will turn into a "summer mini-series", playing out in a string of potentially embarrassing hearings.

That could stifle any effort to increase the NIH's 2005 budget above the 2.6% increase proposed for the agency by the president George W. Bush in February. NIH supporters may plead with law-makers to give the agency an 8–10% budget increase, but they will get nowhere fast, the lobbyists say. "This issue gives a handy excuse to folks in Congress who are saying: we just doubled your budget, how can you come back asking for more?" says one experienced NIH supporter. "This adds to what is shaping up to be a very bad environment for NIH advocacy." ■

Software company bans competitive users

Jim Giles, London

Chemists who have been banned from using a leading software package because they use competing products have taken their fight online.

The group has set up an anonymous website to draw attention to the behaviour of Gaussian of Wallingford, Connecticut.

The site, which appeared about three months ago, provides a list of scientists who are “banned by Gaussian”, and says it aims to “shed light on some practices of Gaussian, Inc. that can undermine basic scientific ideals”.

Gaussian’s software is used by theoretical chemists around the world to predict molecular properties such as bond length and reaction energy. Thousands of site licences for Gaussian have been sold at about US\$2,000 each for academic use.

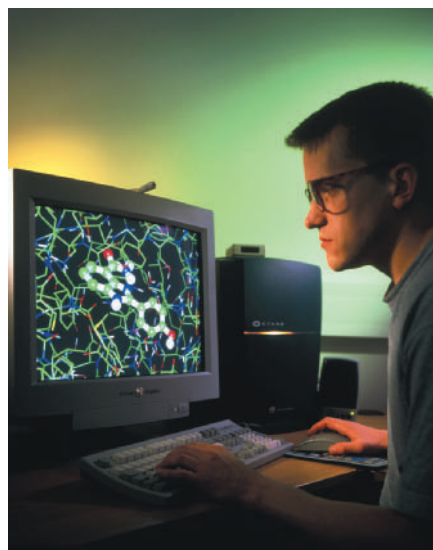
But a small number of prominent researchers have been banned from using the software. Since the early 1990s, Gaussian has denied licences to chemists who work on software that replicates parts of Gaussian’s functionality.

Researchers who are issued with licences must also agree not to allow access to banned chemists. At least ten scientists have been denied licences.

“It’s like a worker from Ford not being able to buy a Chrysler car,” says Peter Gill, a banned chemist from the University of Nottingham, UK.

The banned researchers say that the company’s actions go beyond normal competitive practices. They argue that the restrictions prevent them from acting as referees on other papers, and from properly checking PhD theses, as they are unable to use the software to replicate their colleagues’ results.

“I have only reviewed about half of the 40 to 50 papers I normally review in a year,” says Mark Gordon, a theoretical chemist at Iowa



Log off: Gaussian fears that rivals could use its programs to “catch up with less effort”.

State University. Gordon found out he was banned when he tried to buy Gaussian software about 18 months ago. He works on a theoretical-chemistry package known as GAMESS, which is distributed for free.

Gaussian’s president, Michael Frisch, says that providing licences to competitors would allow rivals to learn about Gaussian’s features without having to develop those features themselves. “Having a competitor’s program allows them to catch up with less effort,” he says. He also says that some researchers with competing products have previously published inaccurate comparisons of the speed of their programs compared with Gaussian.

Frisch acknowledges that it is good practice to check calculations during peer review, but denies that it is essential.

“Provided that the method is written up and published I don’t normally redo calculations,” agrees Mike Robb, a theoretical chemist at Imperial College London, who has worked on the development of Gaussian.

Gaussian says it has no plans to change its policy. Some researchers point out that the problems faced by banned researchers could be reduced if the many individual programs that replicate parts of Gaussian were integrated into a single product.

Researchers funded by the US Department of Energy are working on a software package that combines some of these programs. They declined to be interviewed. ■

Charging plan casts a shadow on Japan’s light beam

David Cyranoski, Tokyo

At the moment, it’s free — but academic researchers may soon have to pay to use one of the world’s most powerful synchrotrons.

Under a plan being considered by the Japanese science ministry, academics would have to join their industrial colleagues in paying to use SPring-8 in Hyogo.

The proposal has already led SPring-8 researchers to begin a worldwide appeal to combat the user fee, which they will say will undermine research priorities at the facility.

SPring-8 has 48 beamlines of intense, stable light that are used by physicists, chemists and biologists to study the structure of molecules and crystalline materials. Like most such facilities, it charges some industrial scientists for beam use. But last month the science ministry ordered the facility to conduct a feasibility study on charging all of its users.

“SPring-8 was expensive to build and has high maintenance costs,” says Yasunori Kojima, director of the ministry’s office of

synchrotron radiation research. “These are tough times, and we have to think about how to make scientists bear some of the burden.”

It is not clear how much the user fee would be. But Akito Kakizaki, a physicist at the University of Tokyo’s Institute for Solid State Physics, is heading a group of 13 researchers who say that any amount would change research priorities for the worse. The group, which is a subcommittee of the Japan Society of Nuclear and Radiochemical Sciences, has written to international synchrotron facilities for support.

The user fees would have to come from grants, which in turn would be beefed up to account for the extra costs. But according to Murray Gibson, director of the Advanced Photon Source in Argonne, Illinois, this system leaves funding in the hands of people too distant from the synchrotron facility to make good decisions.

Gibson has written to Kakizaki saying that he is “seriously concerned” about the

“counter-productive” user-fee proposal.

Michael Chesters, who heads the Daresbury Synchrotron Radiation Source in Warrington, UK, also supports Kakizaki. His facility charged scientists for beam time until two years ago, when the pay scheme was cancelled. The charges made some worthwhile projects look disproportionately costly, says Chesters, leading to inappropriate decisions about who could use the beam. “Once the country has decided it is important to build the facility, it shouldn’t punish those who want to use it most extensively,” he says.

Kakizaki’s group will send its report to SPring-8 officials by the end of this month. A second group, headed by Hiromichi Kamitsubo, a former director of SPring-8, is also working on a report that will evaluate the proposal and the changes in funding needed to make the system work. The science ministry will make its decision about user fees by the end of this year. ■

▶ www.spring8.or.jp/e

Germany backs genome networks to tackle diseases

Quirin Schiermeier, Munich

Germany has confirmed funding for more than 300 post-genomic research projects aimed at finding therapeutic targets for common diseases, including cancer and neurological disorders.

Germany's National Genome Research Network (NGFN) will receive €135 million (US\$163 million) from the science ministry for its second phase, which runs from 2004 to 2007. This includes funding of six large disease-related 'genome networks', each involving molecular geneticists and clinical researchers, and new genomics infrastructures, including a 'mouse clinic' for analysing mutant mice and screening animal models of human diseases.

But some basic researchers are unhappy with the programme's insistence that they collaborate with clinicians, as genomics-based drugs are not even on the horizon. They claim that collaborations involve too much bureaucracy.

"All this networking is really an obsession of politicians," a neuroscientist at one of Germany's leading research hospitals told *Nature*. "You end up feigning how well you collaborate with everybody, just to get funded. But it is actually disadvantaging the best groups."

Other researchers support the approach. "Germany's genome initiative has greatly improved collaboration between clinicians and research facilities," says Martin Hrabé de Angelis, who runs the German Mouse Clinic at the GSF national research centre in Munich.

Research will now focus on identifying candidate genes that might be linked to complex genetic disorders such as cancer or Alzheimer's disease. €10 million has been earmarked for small, high-risk projects.

Announcing the funding, research minister Edelgard Bulmahn pledged that significant funding for genome research will continue until at least 2010.

"Programme-oriented research is the current buzzword," says Friedrich Luft, head of nephrology and hypertension at Charité, the medical faculty at Humboldt University in Berlin. "My intellectual activities are forced into 'networks'. If I do not participate, my chances for funding are greatly diminished, if not zero."

Hrabé de Angelis dismisses such gripes as "small-minded". "Research à la 'one gene, one postdoc' also still exists," he says. "But in a bigger context it is necessary to work together."



Fast track to merger? In lieu of cash, Russia could give ESA increased access to its Soyuz rockets.

Talks pave way for docking between Russia and ESA

Tony Reichhardt

Russian and European space officials are discussing the possibility of Russia joining the 15-member European Space Agency (ESA) — a move that would put ESA on a par with NASA in terms of spaceflight experience and access to orbit.

Although no formal offer has been tendered, a flurry of meetings and public statements in the past few months suggest that both sides are exploring the idea. At the Berlin Air Show last week, the new head of the Russian Aviation and Space Agency (PKA), Anatoly Perminov, told the Russian news agency Itar-Tass that he supports a merger in principle, but that Russia would want full membership of ESA. Perminov, who formerly headed the Russian military space programme, said he has been discussing the details with Jean-Jacques Dordain, who took over as director-general of ESA last July.

Dordain proposed, in a document called Agenda 2007 released last October, that Russia be admitted as an associate member of ESA, with a status similar to that of Canada. Associate members can participate in some agency projects and decisions, but do not have to adhere fully to agency rules. These include the 'juste retour' policy, whereby members receive ESA contracts in direct proportion to how much money they put into agency projects.

Money is precisely Russia's problem.

Although the country has extensive experience in human space flight and some of the most reliable rockets in the world, it is chronically short of funding, and might have difficulty paying its contribution as a full ESA member. But "where there's a will, there's a way", says Giovanni Bignami, director of CESR, the laboratory of space astrophysics in Toulouse, France, and chair of ESA's Space Science and Advisory Committee. ESA, for example, could rewrite its rules to allow member states to pay in kind, with services instead of cash.

ESA has been building steadily closer ties with Russia, and is constructing a launch pad at its site in Kourou, French Guiana, for the Russian Soyuz rocket, with launches to begin in 2007. The European agency currently has 15 member states, with Greece and Luxembourg signed on to become full members by the end of next year.

If the *juste retour* policy and other details can be worked out so that Russia can join, Bignami says that it would be "an enormous success for

ESA". And Russia would get a steady customer for its space industry. Cooperation with the United States has been complicated in recent years by US national security regulations, which include a law that prohibits NASA from buying Russian space hardware outright, because of concerns that Russia is supplying nuclear technology to Iran.



Jean-Jacques Dordain: holding talks with Russia.



Keen: Anatoly Perminov wants full membership.

Big buzz as cicadas arrive after 17-year gap

Laura Nelson, Washington

The emergence of periodical cicadas across the eastern United States, after 17 years underground, is offering scientists an infrequent opportunity to study the insects' interaction with the environment.

Trillions of the *Magicicada* will appear this month to fornicate furiously for about six weeks and then die, while a rather smaller number of entomologists, evolutionary biologists and ecologists grab the chance to study the phenomenon. Newly hatched cicadas will then migrate underground and feed on tree roots for another 17 years.

The regular timing of the emergence intrigues evolutionary biologists. Other researchers will study how human behaviour affects the cicada population, and how the intense insect swarm interacts with its habitat.

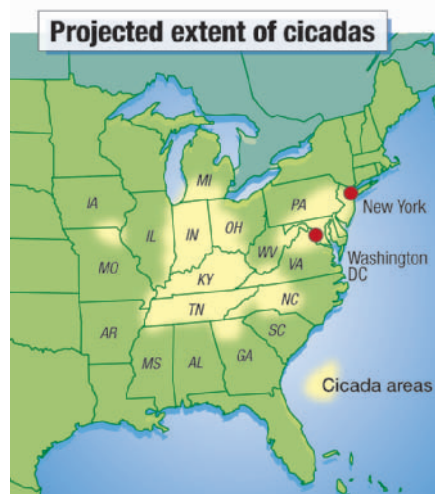
"This is a 17-year experiment," says Keith Clay, a biologist at Indiana University, Bloomington, and one of about a dozen senior US scientists whose professional lives are devoted to the study of periodical cicadas. For them, this spring's buzz is the culmination of years of preparation.

Indiana has wide swathes of deciduous forest and is a "vast natural laboratory," says Clay. To investigate the impact of cicadas on trees, he has created a chessboard of patches of plastic sheeting at a military base in Crane, Indiana, to compare the condition of trees before and after the insects emerge from the ground.

Despite the fears of some, who say that massive urbanization on the eastern seaboard will disrupt the cicadas' cycle, most researchers expect a bumper crop of the insects this year. Reforestation of disused farmland in the region, together with suburban sprawl,



Tree for all: trillions of cicadas will emerge to lay their eggs over the next six weeks.



should facilitate their breeding.

"We have more forest now than at any time in the past 100 years," says John Odland, a geographer at Indiana University. Cicadas favour young trees that will last for many generations, as well as the warm sunlit areas around the edges of forests. Such conditions are common in the suburbs. Researchers say that in some places there will be more than 370 noisy, colourful insects for each square metre. The scene will be "like a science fiction movie", enthuses May Berenbaum, who is an entomologist at the University of Illinois at Urbana-Champaign.

James Speer, a biogeographer at Indiana State University, Terre Haute, is studying the cicadas' impact on forestry by analysing tree-ring formation to check growth rates in cicada-intensive regions. Growth is hardest hit just before emergence, he explains, when cicadas are guzzling roots in preparation for their six weeks in the sun. "After emergence, tree growth rate increases again," he says.

But perhaps the most interesting question about the creatures is why they spend 17 years underground, a cycle that has been recorded since the mid-nineteenth century.

Biologists think it is no coincidence that cicadas have evolved to reappear with periods that are large prime numbers. If the life cycle of a breed was 12 years, they speculate, it would interbreed frequently with others that had life cycles of 2, 3, 4 or 6 years. "The prime number prevents mating between two broods and hybridization," says Christine Simon, an evolutionary biologist at the University of Connecticut. There are 13-year and 17-year broods, and these can only meet up every 221 years, giving them ample time to forge their unmistakable identities.

www.cicadamania.com

Britain plans laws to restrain animal-rights activists

Jim Giles, London

The British government has committed itself to legislation that will specifically target crimes committed by animal-rights activists. But early indications are that the law will fail to satisfy research lobby groups.

Calls for such legislation have intensified following a decision in January by the University of Cambridge to scrap ambitious plans for a primate research centre. The plan was abandoned in part because of the extra security costs of protecting the facility and its staff against attacks by activist groups, some of which have gained notoriety for their violent attacks on people and property associated with animal research (see *Nature* 428, 882; 2004).

New laws designed to curb such protests are now being developed, science minister David Sainsbury told the House of Commons Science and Technology Committee on 12 May. Sainsbury revealed little about the legislation — but he did say that it will be tacked on to another bill.

But groups such as the Research Defence Society, a lobby group that defends animal testing, say that a bill is needed that is dedicated fully to animal research, comparable to the Football (Disorder) Act 2000, which was introduced to tackle violence at soccer matches. Lobbyists say this would ensure that tough penalties can be imposed for crimes committed in the name of animal rights. "The government has

added bits to bills before and it hasn't solved the problem," says Mark Matfield, director of the society.

Sainsbury did earn praise from Matfield and others for confirming that the legislation will address one particular cause for concern — protests outside the homes of staff of companies with links to animal research.

The Home Office, which will draft the new legislation, says it is too early to say what form these restrictions will take. A spokeswoman says that a previous set of laws designed to tackle animal-rights extremists was implemented in January, and needs to be monitored before further measures can be finalized.

US changes rules to speed up approval of AIDS drug cocktails

Geneva The United States is to fast-track regulatory approval of a group of AIDS drugs tailored to meet the needs of developing countries.

Combination drugs — single pills containing cocktails of previously approved substances — simplify distribution and make treatment courses easier to follow in areas with poor health infrastructure. Until now, approval took six months or longer, but US Health and Human Services secretary Tommy Thompson said on 16 May that the process will in future be complete in two to six weeks.

As the constituent drugs have already passed regulatory approval, the US Food and Drug Administration will not require new trials of a combined pill, provided the data shows that mixing the drugs does not interfere with their chemistry or pharmacology. The scheme will also be open to companies selling branded drugs and foreign producers of generic versions of AIDS drugs.

The decision was announced on the eve of the World Health Organization's annual assembly in Geneva, Switzerland.



What's your poison? Puffer fish usually contain tetrodotoxin, which can be fatal to diners.

Non-toxic puffer fish takes the 'die' out of dining

Tokyo Gourmets in Japan may soon be able to feast safely on puffer-fish liver — a delicacy currently banned because of the toxins it contains.

The dish has been illegal since 1983; before that hundreds of people died every year after eating it. But Tamao Noguchi, a fisheries scientist at Nagasaki University, is trying to get it back on the menu.

Puffer fish (*Takifugu*) are believed to accumulate poison by eating starfish and other toxic creatures that live at the bottom of the sea. To produce toxin-free fish, Noguchi's group raised about 5,000 of them

in nets suspended ten metres above the sea floor or in tanks with purified sea water. On 13 May at a meeting of the Food Hygienics Society of Japan in Tokyo, the researchers reported that all the fish were poison-free.

The group has patented the technique and hopes to create a special district where people can feast on the plump organs. Team member Osamu Arakawa denied that the lack of the poison, which attracts some people with its lip-numbing effect, would drive away consumers. "It's delicious even without the tingling effect," he says.

CERN supports drive for free access to data

Geneva One of the world's main physics laboratories has thrown its weight behind the open-access movement, the international effort to promote unrestricted access to scientific knowledge.

CERN, the European particle-physics laboratory near Geneva, Switzerland, signed a declaration on 13 May committing the lab to making its research results freely available over the Internet. The declaration, launched last October in Berlin (see *Nature* **425**, 752; 2003), has 38 signatories so far, including all the major research organizations in Germany and the CNRS, France's main national research agency.

CERN already makes papers available online for free, copyright permitting. But its formal support for open access is still significant, says Jürgen Renn, a director at the Max Planck Institute for the History of Science in Berlin and one of the declaration's designers.

Panther's land clawed back as fur flies over 'flawed' data

San Diego The US Fish and Wildlife Service has been accused of using flawed scientific data to reduce the habitat area of the endangered Florida panther (*Puma concolor coryi*) and facilitate property development.

The complaint, filed on 4 May by Andrew Eller, a biologist at the wildlife service, alleges that senior administrators manipulated data about the panther population and habitat needs to reduce restrictions on private land slated for development east of Naples, Florida. The land is near wildlife parks that are home to the remaining 87 Florida panthers.

Eller, who is backed by Public Employees for Environmental Responsibility, a Washington-based alliance of government agency staff, has worked for ten years on panther recovery efforts in Florida.

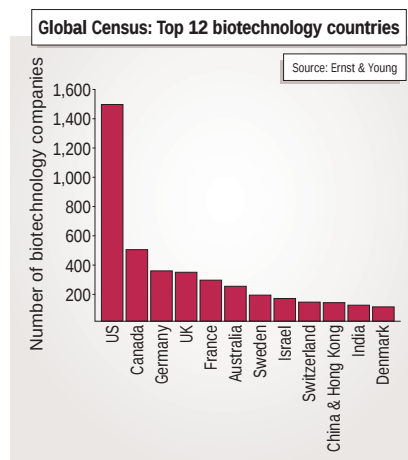
The service has 60 days to respond to the complaint. Jay Slack, a field supervisor with

the service's Vero Beach office in Florida, denied the accusations. The agency "has led the fight," Slack says, to "protect both the species and its habitat".

Europe's biotech firms hit by cash shortfall

Munich European biotechnology companies have dangerously thin reserves, according to Ernst and Young's 2004 Global Biotechnology reports.

The number of listed biotech firms in Europe continues to increase, the authors



say, but 43% of publicly traded companies have less than two years' worth of cash. And although total investment has more than doubled to US\$2.6 billion during 2003, it still falls far short of the \$14.4 billion raised by US firms last year. The report blames bureaucracy and the slow pace of European Union legislative reform for hampering the sector's development.

The number of US companies making money could rise, the report adds. After a period of severe cash shortage caused by stock-market falls, the US biotech industry, which has drugs awaiting approval, is likely to become profitable by 2008, it claims.

Howard Hughes bumps up number of investigators

Washington The US Howard Hughes Medical Institute is to expand the roster of scientists on its payroll.

The \$10-billion philanthropic organization said on 13 May that it would recruit 30 to 50 new investigators for 2005, on top of the 318 researchers it already supports. The grants are much sought after, as they cover all research costs for around five to seven years at a time.

The institute will also award \$50 million in new grants for undergraduate and masters' programmes.

Not so special, after all?

The parasite *Giardia* was thought to represent a throwback to the earliest days of advanced cellular life. But biologists are now arguing over its true status. Jonathan Knight reports.

Intestinal cramps rarely inspire deep thoughts about evolution. But anyone who has suffered an infection of the gut parasite *Giardia intestinalis* can boast of having an intimate, if painful, familiarity with an icon of evolutionary theory. Textbooks have for years held up this single-celled parasite as a living relic of the very early days in the evolution of plants, animals and other higher life forms.

But this may have to change. A new discovery about *Giardia*'s own innards, which makes the parasite look rather more run-of-the-mill, has caused a rift among evolutionary biologists. Some argue that its role in the history of life will have to be reconsidered. But others are fighting to retain *Giardia*'s iconic status.

At the heart of this debate lie mitochondria, sacs of enzymes that use oxygen to generate energy for cells. Organisms from people, through house flies and elm trees, to single-celled yeast and algae, all have them. These are the eukaryotes, whose cells contain many types of other useful sacs, or 'organelles', as well as a nucleus, all packaged in the same sort of fatty membrane that surrounds the cell.

The rest of life lacks this equipment. Prokaryotes, which include bacteria and archaea — ancient, single-celled creatures found in extreme environments such as volcanic springs — are devoid of mitochondria and other specialized organelles.

Biologists now believe that eukaryotes gained their mitochondria sometime between 1.5 billion and 2.2 billion years ago¹. At the time, rising oxygen levels in Earth's atmosphere were causing problems for organisms adapted to life without the gas. The theory goes that the ancestor of mitochondria was a prokaryotic cellular parasite



A taste of the past: the gut parasite *Giardia intestinalis* (green) has for years been an evolutionary icon.

that could process oxygen. It took up residence inside an ancestral eukaryote and helped it by soaking up toxic dissolved oxygen. Over time, this parasite became an indispensable source of chemical energy for the host, using oxygen to 'burn' sugars. Most of its genes relocated to the host's nucleus, sealing the partnership.

Primitive appearance

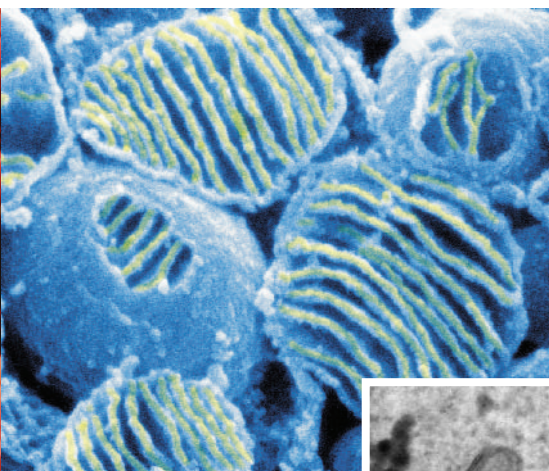
Until a few months ago, *Giardia* was thought to represent a throwback to the time before this union. Mitochondria are relatively easy to see under the microscope because they are large and have a characteristic appearance. But none were ever seen in *Giardia*, so it was thought to be a direct descendent of pre-mitochondrial eukaryotes. Partly as a result, resources have been ploughed into *Giardia* research, including a project to sequence its genome. "A lot of people's efforts were invested in the idea that *Giardia* was the most primitive eukaryote," says Andrew Roger, who studies molecular evolution at Dalhousie University in Nova Scotia, Canada.

Doubts about *Giardia*'s evolutionary position first began to emerge in 1998, when Roger and his colleagues found a strong similarity

between a *Giardia* gene, called *cpn60*, and a gene from the α -proteobacteria². These bacteria, which include the pathogen that causes typhus, are thought to be the closest living relatives to the prokaryotes that gave rise to mitochondria. Given the genetic similarity, Roger and his colleagues concluded that *Giardia* once had mitochondria but had lost them.

But the gene might also have been acquired from another invader that took up residence in eukaryotic cells and did not evolve into mitochondria. This is not as improbable as it might sound — hundreds of examples of such 'endosymbiosis' are known³. So far as we know, in billions of years, endosymbiosis has led to the formation of only two cellular organelles: mitochondria and chloroplasts, the site of photosynthesis in plants and algae, which are believed to be related to modern cyanobacteria.

But the real bombshell came last November, thanks to an international team led by molecular parasitologist Jorge Tovar of Royal Holloway, a college of the University of London in Egham, Surrey. His team found in *Giardia* a number of proteins related to those associated with mitochondria in other eukaryotes. More significantly, these



Body of evidence: it was thought that *Giardia* had never evolved mitochondria (above), but electron micrographs (inset) have now revealed it contains sacs that resemble diminished versions of this cellular equipment.



proteins were clustered together at discrete locations in the organism's cells. Using electron microscopy, the researchers saw tiny sacs at those spots, which they identified as mitochondrial relics called mitosomes⁴.

Gut feeling

Tovar first identified mitosomes in 1999 in an unrelated organism called *Entamoeba histolytica*, a deadly intestinal parasite⁵. Like *Giardia*, *Entamoeba* was once considered to have branched off the trunk of the eukaryotic evolutionary tree before mitochondria came on the scene, and the two were lumped together in a class called Archezoa⁶. But in the 1980s, DNA sequence evidence showed that *Entamoeba* branched much later than

other organisms with fully fledged mitochondria⁷. To many researchers, this suggested that it had lost its mitochondria at some point. Tovar's discovery of mitosomes indicated that they were not entirely gone, only much reduced.

For Tovar and many others, the evidence that *Giardia* once had mitochondria is solid. For example, the electron micrographs of its mitosomes revealed a double membrane surrounding the sacs. This is a property of only three known cellular structures: chloroplasts, the nucleus and mitochondria. Furthermore, contained within those membranes are two proteins that process iron-sulphur clusters, complex structures that are essential to the mitochondrial job of energy production. "We now know that *Giardia* diverged after the evolution of mitochondria," Tovar asserts.

But Patricia Johnson, an evolutionary biologist at the University of California, Los Angeles, is among those who are not convinced. Johnson studies the evolution of hydrogenosomes, another cellular sac that may similarly be the product of mitochondrial reduction. One problem, she says, is that the protein produced by *cpn60* does not migrate to Tovar's mitosomes. Yet in fully fledged mitochondria and in other microorganisms with mitosomes or hydrogenosomes, it does. "The data make it a toss-up, but it's not being treated with that kind of caution," she says. To strengthen his case, Tovar is trying to identify all the proteins in *Giardia* mitosomes, to see if more of those normally associated with mitochondria are present.

Mitchell Sogin of the Marine Biological Laboratory in Woods Hole, Massachusetts, who heads the *Giardia* genome project, also wants to see more evidence before changing his mind about the parasite's evolutionary status. He points out that sequence analysis shows *Giardia*'s iron-sulphur cluster genes to be only weakly related to mitochondrial genes⁸. "Evolutionary trees must be strong if you are making strong statements about evolutionary history," he says.

This attitude frustrates people such as William Martin, who studies molecular evolution at Heinrich Heine University in Düsseldorf, Germany. He is convinced that the best and simplest explanation for the data is that *Giardia* once had mitochondria⁹. Some people, he argues, refuse to accept this because they have spent too many years working on the opposite

assumption. "They don't want it to have mitochondria because it spoils their soup," he says. "This thinking is deeply ingrained."

The thinking has its roots in the concept of the Archezoa, Martin argues, the group that was conceived to bring together a range of single-celled eukaryotes thought to lack mitochondria. *Giardia* was the granddaddy, having branched off on its own before any other eukaryote, according to evolutionary trees built using sequences of RNA from ribosomes, the organelles in which proteins are made (see Figure, below left).

Family ties

But one by one, the Archezoa all proved to have either a set of mitochondrial genes in their nuclei, or relics of mitochondria such as mitosomes or hydrogenosomes. This isn't entirely surprising, because many of these organisms are parasites, which tend over evolutionary time to become stripped-down versions of their former selves as they become increasingly dependent on their hosts.

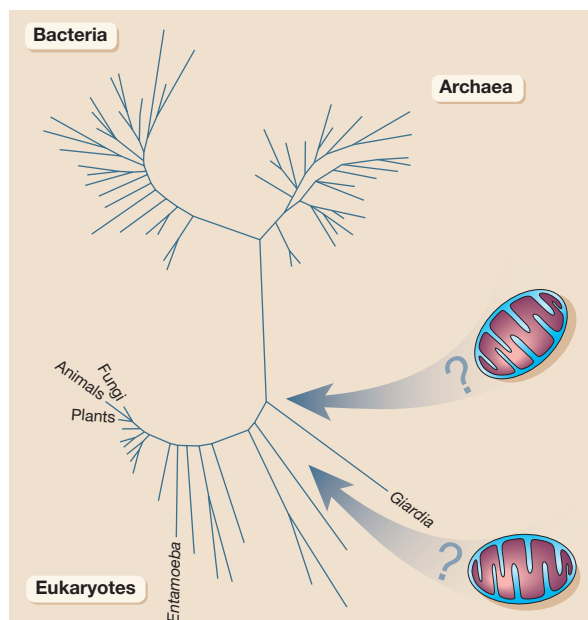
Giardia's status as the earliest branching eukaryote has also been questioned. Some recent evolutionary trees that take into account the variable rates at which different DNA bases mutate paint a much muddier picture of the early branches¹⁰.

If Tovar, Roger and Martin are correct, how bad is the news for *Giardia* research? Sogin contends that even if *Giardia* is shown definitively to have once had mitochondria, it would have no adverse impact on his genome project. Apart from the organism's importance to human health, its sequence should still reveal much about evolution because it branched off the eukaryotic tree very early, he says.

But *Giardia*'s status as a uniquely important throwback to the earliest days of the eukaryotes surely would be usurped if someone discovers a new member of the Archezoa, sans mitochondria or mitosomes, lurking in the oxygen-starved muck at the bottom of a lake. This possibility hasn't escaped the notice of Roger and others, who are still collecting samples in the hope of finding something new. "There is a lot more out there, and it's possible that such organisms still exist," Roger says.

Jonathan Knight writes for Nature from San Francisco.

1. Dyall, S. D., Brown, M. T. & Johnson, P. J. *Science* **304**, 253–257 (2004).
2. Roger, A. J. et al. *Proc. Natl. Acad. Sci. USA* **95**, 229–234 (1998).
3. Buchner, P. *Endosymbiosis of Animals with Plant Microorganisms* (Interscience, New York, 1965).
4. Tovar, J. et al. *Nature* **426**, 172–176 (2003).
5. Tovar, J., Fischer, A. & Clark, C. G. *Mol. Microbiol.* **32**, 1013–1021 (1999).
6. Cavalier-Smith, T. in *Endocytobiology II* (eds Schwemmler, W. & Schenk, H. E. A.) 265–279 (de Gruyter, Berlin, 1983).
7. Sogin, M. L., Edman, U. & Elwood, H. in *The Hierarchy of Life* (eds Fernholm, B., Bremer, K. & Jörnvall, H.) 133–143 (Elsevier Science, Amsterdam, 1989).
8. Tachezy, J., Sánchez, L. B. & Müller, M. *Mol. Biol. Evol.* **18**, 1919–1928 (2001).
9. Henze, K. & Martin, W. *Nature* **426**, 127–128 (2003).
10. Simpson, A. G. B. & Roger, A. J. *Curr. Biol.* **12**, R691–R693 (2002).



A question of timing: did advanced cellular life gain mitochondria after *Giardia* branched off from the evolutionary tree or before?

It's the night shift at a nuclear power plant in upstate New York, and a technician checks on the temperature of the reactor core. It's a searing 800 °C — a temperature that today would spark a major panic and could signal the start of a partial reactor meltdown. Yet this reading doesn't raise an eyebrow. It's 2035, and this state-of-the-art reactor is designed to operate at this temperature, cooled not by the water that keeps today's reactors in check, but by a huge vat of molten lead. And thanks to its high operating temperature, the plant is generating hydrogen fuel as well as electricity.

This vision of the future comes from the Generation IV International Forum (GIF), a consortium of ten nations that is planning the nuclear reactors of tomorrow. These new plants would all operate at high temperatures, improving their efficiency. And they would include simplified safety features that do not rely on sophisticated backup systems or experienced operators — all are, in principle, 'meltdown proof' and can cool themselves down in the event of an accident with minimal, if any, human intervention.

This would also mean that any attempt to trigger an accident deliberately — by shutting off the coolant or power supply — would be in vain. Nuclear reactors would become less of a terrorist target. At least, that's what the nuclear industry hopes.

But given memories of the partial reactor meltdown in 1979 at Three Mile Island, Pennsylvania, and of the 1986 accident at Chernobyl in Ukraine, during which the reactor core exploded, killing some 31 people immediately and spewing radioactive debris across Europe, the public will take some convincing.

No new nuclear power plants have been ordered in the United States since the accident at Three Mile Island. But in March this year, a consortium of US energy companies said that it intends to apply for a licence, the first step towards building a plant. In April, France announced that it would replace its elderly 59 reactors with new ones. And Asian countries are planning to build dozens of reactors to cope with their booming energy demands. Is this the beginning of a nuclear revival?

Fuel for the future?

Nuclear power is not a source of carbon dioxide, and with emissions of this greenhouse gas now soaring, and global energy demands predicted to double by 2050, the nuclear option is finding its way back onto the table. At a 2002 gathering of GIF representatives in Tokyo, Spencer Abraham, the US energy secretary, used these arguments to explain the Bush administration's strong support for nuclear energy. If nuclear engineers can overcome the technical hurdles involved in building the next generation of reactors, Abraham said, then we will have

energy that is "safe, abundant, reliable, inexpensive and proliferation resistant".

For nuclear power to undergo a renaissance, experts agree that reactors will need to be a lot cheaper to run. And to sway a nuclear-averse public, the next generation of reactors will need to produce much less radioactive waste at terrorist-proof facilities.

Such technological challenges are too great for one country alone. In 2001, the eight founding nations of GIF decided to pool their research expertise, and later picked what they believe are the six best prospects for the reactors of the future¹ (see Table, opposite).

No more than three of the six designs are likely to survive the feasibility testing phase and go on to become research prototypes, each costing about US\$1 billion to build and test, predicts William Magwood, director of the US Department of Energy's Office of Nuclear Energy, Science, and Technology, and chairman of GIF's board.

Unlike today's water-cooled reactors, which tend to run at about 300 °C, all six concepts are designed to run at temperatures from 510 °C to 1,000 °C. This allows for more efficient conversion of heat to electricity — one leading design, the very-high-temperature reactor (VHTR), could squeeze 50% more electricity from the same amount of fuel compared with conventional plants.

But these higher operating temperatures mean that the reactors will need new coolants, as ordinary water can only be used, under typical pressurized conditions, up to

330 °C. Two GIF concepts use inert helium to keep the reactor cool; others use molten lead, sodium or salt.

One of the most popular generation IV concepts, the supercritical-water-cooled reactor (SCWR), uses extreme pressures to prevent water from boiling at temperatures up to 500 °C. Because of the reactor's efficiency and relatively simple design, it would potentially be fairly cheap to build and run, says Jacopo Buongiorno, integration manager for the SCWR system at the Idaho National Engineering and Environmental Laboratory in Idaho Falls. Indeed, if it works, it will churn out electricity at prices that, on paper, are competitive with coal and gas, and vastly cheaper than existing nuclear reactors.

Hotting up

In terms of practical experience, perhaps the most advanced concept is the VHTR. Japan's Atomic Energy Research Institute based in Kashiwa already operates a similar high-temperature engineering test reactor at Oarai, near Tokyo. This reactor is cooled by helium gas, and it reached its operating goal of 950 °C for the first time in a test run last month.

At temperatures of about 700–900 °C, reactors can be used to split hydrogen from water thermochemically. Many countries that largely depend on oil for their energy needs are betting on hydrogen as the fuel of the future, using fuel cells to convert the gas into electricity for cars and homes.

Without a switch to hydrogen, the energy

Nuclear power's new dawn

Global warming and rising energy needs are rehabilitating the concept of nuclear power. But if it is to figure in the energy equation, it will need to be cheaper, cleaner and safer, says Declan Butler.

Energy has made the VHTR its number one choice. A US energy bill that would provide \$1.1 billion to build a prototype at the Idaho lab by the middle of the next decade is currently held up in Congress by disputes over unrelated issues, such as energy regulation.

"GIF recognizes nuclear's role in transportation. We've raised the stakes by including demonstration of hydrogen production in some of the reactor designs," says Ralph Bennett, the Idaho lab's director of advanced nuclear energy. But generating enough hydrogen to completely replace gasoline for the United States' transport needs would mean building more than 400 nuclear power plants, each generating a gigawatt³. There are only 441 nuclear plants in the world today.

Few in the industry dispute the wisdom of a shift towards high-temperature reactors, but only two of the designs — the VHTR and the SCWR — would be able to operate without having to depend on the controversial reprocessing of plutonium waste (see 'Plutonium wars', overleaf).

All the designs include untested engineering, and also depend on the development of new ultrahard materials that can resist continued high temperatures, intense bombardment by neutrons in the chain reaction, and often corrosive reagents, says David Lochbaum, a nuclear engineer with the Union of Concerned Scientists, an environmental pressure group based in Cambridge, Massachusetts. In the molten salt reactor, for example, uranium fuel is dissolved in the circulating coolant, and this would need new

corrosion-resistant materials to prevent any possibility of a radioactive leak.

Magwood agrees that developing new materials will be one of GIF's greatest challenges. There are plans within GIF for a bold international research programme in this area that could also develop materials for nuclear fusion reactors, which face many of the same problems. But no decision is expected until summer 2005, when the GIF partners are scheduled to finish thrashing

out plans for who will do what research on each of the six concepts.

Operating at high temperatures also rules out conventional fuel systems, in which uranium pellets are loaded into metal rods, as the rods melt at fairly low temperatures.

Instead, the gas-cooled reactors will hold fuel pellets either in a honeycomb graphite structure, as in the Japanese test reactor, or fused into billiard-ball-sized graphite spheres, known as pebbles.

Core issues

In pebble-bed reactors, millions of these billiard balls are loaded into the core, and gas coolant flows through the spaces to remove the heat. The balls can be continually removed from the bottom of the reactor and are sorted automatically — those that are almost spent are sent to a waste stream, and those with some life left in them are returned to the top of the pile.

Each pebble is itself a mini-nuclear-reactor. A core of fuel is coated with a layer of graphite that slows down neutrons to control the nuclear chain reaction. This is covered with an ultrahard ceramic layer, sealing in all the fission products. In principle, this should

"Without successful generation IV concepts, the nuclear industry will struggle to maintain its current position of generating 17% of the world's electricity."

demand from US transport alone will cause the nation's oil imports to soar by 78% by 2025 (ref. 2), so the US Department of

The six generation IV reactor concepts

The Generation IV International Forum (GIF) consists of the United States, Argentina, Brazil, Canada, France, Japan, South Korea, South Africa, Switzerland and the United Kingdom

	Coolant	Temp. (°C)	Pressure	Waste recycling	Power output	Research needs	Earliest delivery
Gas-cooled fast reactors	Helium	850	High	Yes	Electricity and hydrogen	Irradiation-resistant materials, helium turbine, new fuels, core design, waste recycling	2025
Lead-cooled fast reactors	Lead-bismuth	550–800	Low	Yes	Electricity and hydrogen	Heat-resistant materials, fuels, lead handling, waste recycling	2025
Molten salt reactors	Fluoride salts	700–800	Low	Yes	Electricity and hydrogen	Molten salt chemistry and handling, heat- and corrosion-resistant materials, reprocessing cycle	2025
Sodium-cooled fast reactors	Sodium	550	Low	Yes	Electricity	Safety, cost reduction, hot-fuel fabrication, reprocessing cycle	2015
Supercritical-water-cooled reactors	Water	510–550	Very high	Optional	Electricity	Corrosion and stress, water chemistry, ultrastrong non-brittle materials, safety	2025
Very-high-temperature reactors	Helium	1,000	High	No—waste goes directly to repository	Electricity and hydrogen	Heat-resistant fuels and materials, temperature control in event of accident, high fuel burn-ups	2020

SOURCE: GIF

prevent any radioactive material escaping in the event of an accident.

The Oak Ridge National Laboratory in Tennessee worked on pebbles back in the 1970s, when high-temperature gas reactors were explored as a source of tritium, a hydrogen isotope once used in nuclear weapons. But one out of every thousand fuel pebbles was either chipped or poorly coated — unacceptable defects for a modern reactor.

All of the generation IV designs face similar hurdles. “The basic research gap is massive,” says Alain Bugat, who heads France’s Atomic Energy Commission. “This is long-term research; if we have a working demo of some of the designs by 2030 we will be doing well.”

Without successful generation IV concepts, the nuclear industry will struggle to maintain its current position of generating 17% of the world’s electricity. A study published last year⁴ by a group of scientists and economists at the Massachusetts Institute of Technology (MIT) in Cambridge looked at the potential for the power generated by nuclear energy to triple by 2050, the level needed for it to have a significant impact on predicted carbon dioxide emissions. The group concluded that, economically, new

nuclear plants would not be able to compete with coal and gas. Worldwide, most electricity markets are being deregulated, which means that there are fewer state subsidies to prop up the nuclear industry.

Substantial numbers of new plants will only be built, the MIT study predicted, if their costs can be cut by a quarter from existing designs, or if a hefty carbon tax is imposed on fossil fuels. The latter seems unlikely, so cost cutting is vital.

But the outlook is not quite so gloomy everywhere. “Most of the expected nuclear growth up until 2025 will be in eastern Asia, and in particular in just four countries — China, India, South Korea and Japan,” says Peter Fraser, a nuclear analyst at the

International Energy Agency in Paris. These will account for more than 85% of all new plants built, he predicts.

The reactors that the Asian countries want to build are mostly generation III systems — revamped versions of today’s generation II reactors, but with multiple backup systems to enhance safety, and with simplified, cheaper designs that have fewer parts to go wrong.

Nevertheless, nuclear output worldwide is more likely to shrink until 2025, as older

plants close more frequently than new ones open. France is one of the exceptions. It gets 80% of its electricity from nuclear power, and intends to begin replacing its ageing plants using a €3-billion (US\$3.5-billion) generation III reactor, built by the Franco-German company Framatome ANP, headquartered in Paris.

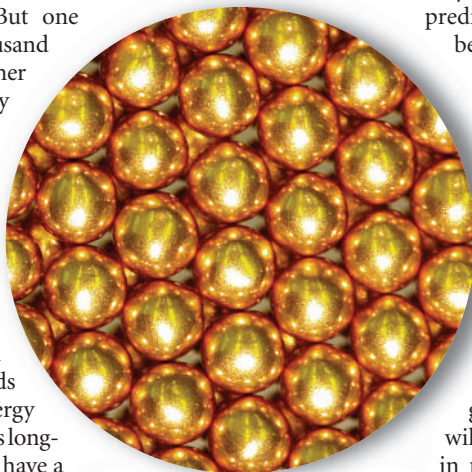
Countries such as France and Japan want to maintain their energy independence, so the decision to rebuild is largely political, not economic. And the economic equation is changing all the time — by 2025 other emerging energy technologies may outcompete nuclear power. “We expect solar energy costs to fall dramatically,” says Fraser.

Such technologies also fit better with the current trend towards decentralized electricity generation in smaller power plants⁵, Fraser notes. In addition, future technologies may help to reduce carbon dioxide emissions from gas and coal plants.

The nuclear industry has to adapt to this rapidly changing global environment. Even if GIF can develop reactors that are supersafe and superclean, unless they are markedly cheaper than competing technologies, the nuclear industry will, for decades to come, be running hard just to stand still. There is likely to be much early morning jogging before any new nuclear dawn.

Declan Butler is Nature’s European correspondent.

1. *A Technology Roadmap for Generation IV Nuclear Energy Systems* (US DOE NERAC/GIF, Washington DC, 2002) online at http://gif.inel.gov/roadmap/pdfs/gen_iv_roadmap.pdf
2. *International Energy Outlook 2004* (US DOE, Washington DC, 2004) online at <http://eeec.eia.doe.gov/oi/af/ieo>
3. Grant, P. M. *Nature* **424**, 129–130 (2003).
4. *The Future of Nuclear Power* (MIT, Cambridge, MA, 2003).
5. *Nature* **427**, 661 (2004).



Great balls of fuel: these pebbles are ‘mini-reactors’ and could be used to power the next generation of nuclear plants.

Plutonium wars

Reprocessing nuclear waste is controversial because it separates out plutonium, a key ingredient of advanced nuclear weapons. To counter plutonium proliferation, the United States, for example, has historically refrained from reprocessing waste, and discouraged other nations from doing so.

But four of the six proposed generation IV reactors would turn this doctrine on its head, as they would require reprocessing, albeit in a form that is supposed to be ‘proliferation resistant’.

Conventional reprocessing recovers plutonium and uranium from waste, and these can be burnt in reactors as mixed oxide fuel (MOX). This process is most-developed in France, which operates a reprocessing facility at La Hague, and at Sellafield in the United Kingdom. Japan, China and Russia also have reprocessing plants.

Reprocessing is attractive because it could cut the final amount of waste produced — 96% of spent fuel consists of uranium and plutonium, whereas troublesome long-lived radionuclides account for less than 1%. So if these long-lived

elements could be extracted from spent fuel, all of the uranium and plutonium could be recovered and reused. Only a small volume of spent fuel would be left over as waste.

But conventional reprocessing not only generates weapons-grade plutonium, it is expensive. Reactors also have to be substantially modified to burn MOX, and MOX itself can only be partially burnt, meaning that a lot of the fuel still ends up as waste.

Supporters of the generation IV reactors claim that new forms of reprocessing can be developed that avoid these drawbacks. For example, additional steps in the process could convert the long-lived waste into elements with shorter half-lives, slashing the time this waste needs to be stored from more than 300,000 years to just centuries.

And the plutonium generated could be spiked with heat-producing and radioactive elements to make it too hot to handle, they argue. Unlike a canister of pure plutonium, which can be picked up safely with a pair of thick gloves, such material would be harder to steal or use.

But the latest techniques being tested in the United States fall far short of these goals, says Edwin Lyman, a senior scientist with the Union of Concerned Scientists, an environmental pressure group based in Cambridge, Massachusetts.

“Weapon-usable plutonium could be extracted by turning a dial, or at most, adding a clean-up stage to a reprocessing plant,” agrees Frank von Hippel, a nuclear physicist at Princeton University, New Jersey, and a former assistant director for national security in the White House Office of Science and Technology Policy.

Ralph Bennett, director of advanced nuclear energy at the Idaho National Engineering and Environmental Laboratory in Idaho Falls, counters that research on reprocessing is still at a very early stage.

But critics argue that any international research effort in advanced reprocessing would itself spread expertise in the chemistry and metallurgy of radioactive elements, including plutonium. The world should be seeking to eliminate existing stocks of plutonium instead of developing its use, they say.

Bioinformatics: smartest software is still just a tool

The proliferation of ever-more-sophisticated biological software can lead to misuse.

Sir—A recent Letter to *Nature* by R. D. King and colleagues, describing a ‘robot scientist’ (*Nature* 427, 247–252; 2004), gives us pause for reflection not only on our own humanity, but also on the increasing use of software in biology. A quick scan of publicly available Internet databases provides some idea of the hundreds of software programs, and their manifest applications, available for biology. Unfortunately, this very prevalence of software in biology may be creating a problem.

My own observations of colleagues and students suggest that the difficulties they experience using biological software could lead to misuse. They often only use the ‘default’ parameters, either because they don’t know how to adjust the parameters to refine their analysis, or because they are concerned that such changes may affect their results. This gap in knowledge and skill is understandable, given that it is difficult enough to keep up with advances

in biological science — never mind all the details of computer software. Moreover, there is increasing pressure on the biologist to manage, and analyse, a deluge of biological data.

Unfortunately, it seems that, while relying on software as a biological tool, we are not giving it the same careful, and critical, consideration as other tools we use. In particular, we seem to be selectively ignorant, not only of the need to calibrate, characterize and standardize software, but also of how software works.

We routinely calibrate, characterize and standardize our normal laboratory tools (such as pipettes, and chemical or biological assays) and understand how they work. We do this so that we can be confident in the results of the experiment, and our interpretation of the results.

Why should software be any different? The complexity of the software with its millions of lines of code, the underlying algorithm and the mathematics may be too

daunting for most biologists to fully understand. But we should not be so impressed by the perceived efficacy of software that we are willing to forgo the strict scientific procedures and quality control that we apply to other laboratory tools.

The promise of software in biology is exciting and could lead to greater understanding of biological systems. Informatics is undoubtedly changing how biological science is being done. Consequently, biological software is increasing in complexity and sophistication, which is all the more reason to be stringent in its application and in understanding its limitations. Software is a useful and valuable tool, but it should be treated no differently from any other tool in biological science.

Steven Bottomley

School of Biomedical Sciences,
Curtin University of Technology, Kent Street,
Bentley, Western Australia 6102, Australia

Bioinformatics needs a software archive

Sir—In February 2003 the US National Institutes of Health (NIH) released a statement on its data-sharing policy, in which it strengthened its commitment to free exchange of final research data produced by public funds (see *Nature* 421, 877–878; 2003). Little attention has been given, however, to a related question — sharing and protecting the products (software and tools) of bioinformatics research, especially infrastructure generated to support large projects.

As projects and databases evolve, some will inevitably lose funding or be shut down. What happens to the bioinformatics software? In one such case, the Genome Database, an international collaboration set up in support of the Human Genome Project, lost US federal funding in 1998, although it was later reopened with private funding. During the shut-down process, a significant portion of the source code (though not data) was irreparably lost. This database represented an investment of more than US\$50 million by US, European and Asian sources. The loss of the code, a public asset, occurred because there was little supervision during decommissioning.

As a minimum safeguard, we propose the creation of a Bioinformatics Software Archive, in which an archival copy of

bioinformatics software would be maintained in a secure central repository supported by public funding. There are moves in the bioinformatics community to adopt open-source standards for software, which provide a means of online archiving, but these are not suitable for every type of software. As with software released under open-source agreements, a central archive will help stop too many researchers trying to reinvent the wheel (thereby saving research funds).

If funding agencies such as the NIH are serious about data sharing, they should be serious about protecting the software used to produce those data.

A. Jamie Cuticchia, Gregg W. Silk

Genome Database, RTI International, Research
Triangle Park, North Carolina 27709, USA

Autistic geniuses? We’re too ready to pathologize

Sir—While I am in sympathy with the thrust of Allan Snyder’s thoughts on ‘autistic genius’ (*Nature* 428, 470–471; 2004), and the book by Michael Fitzgerald, *Autism and Creativity*, which he reviews, I think there is a danger of going overboard on the subject of pathology and creativity. Snyder cites me as saying that the philosopher Ludwig Wittgenstein had autistic traits, but I have never thought or suggested this. Snyder’s misapprehension

arose from a conversation that I had with the autistic scientist Temple Grandin — reported in my book *An Anthropologist on Mars* (Knopf, New York, 1995) — about another researcher’s suggestion that Wittgenstein may have been autistic. Grandin erroneously attributed this opinion to me in her otherwise excellent book, *Thinking in Pictures* (Doubleday, New York, 1995).

I think that pathologizing genius, and pathologizing historical figures, has become an obsession with us; and also that the concepts of ‘autism’ (and ‘Tourette’s syndrome’, etc.) have become so distended that they are vastly overused. It seems to me extremely unlikely, from the evidence we have, that Wittgenstein or Einstein or Newton were significantly autistic.

On the other hand, there is strong historical evidence to suggest that the eighteenth-century chemist Henry Cavendish was autistic (see O. W. Sacks, *Neurology* 57, 1347; 2001). Unlike most other supposed ‘autistic geniuses’, he showed a near-total incomprehension of common human behaviours, social relationships, states of mind, and money, as well as an almost obsessed attention to detail — which led to the great generalizations he was later to erect. But it is important to sift through the evidence with great care before diagnosing or pathologizing.

Oliver Sacks

2 Horatio Street, New York, New York 10014, USA

Time waits for no man

Do you know what you think before you see what you say?

Mind Time: The Temporal Factor in Consciousness

by Benjamin Libet

Harvard University Press: 2004. 288 pp.

\$29.95, £19.95, €27.70

Kevan Martin

"I did not know I loved you till I heard myself telling you so — for one instant I thought, 'Good God, what have I said?' and then I knew it was the truth." Bertrand Russell's report of his late-night conversation with Ottoline Morrell captures the essential thesis of Benjamin Libet's monograph on the temporal factor in consciousness. Libet reviews his extended set of experiments, which he believes show that Russell was right: reality is running about half-a-second ahead of our conscious awareness of it. The bad news is that we never catch up.

Studies of awareness and consciousness by neurobiologists and psychologists are now mainstream, and even young scientists can expect to pursue a well-funded career in consciousness studies. It was not always so. As Francis Crick records in his own book on consciousness, *The Astonishing Hypothesis* (Simon & Schuster, 1994), Libet only dared switch to the study of consciousness after he got tenure. It is fortunate for us that he did, and that he has presented us here with what amounts to a retrospective exhibition of his work. The book is something of an atavism, resembling more an early work of his mentor Jack Eccles than the glossy volumes that now appear in the field. But the refreshing result is that we are immediately engaged in an earnest one-to-one tutorial with Libet.

Libet was fortunate to link up with the neurosurgeon Bertram Feinstein, who gave him the chance to study the neocortex of conscious patients who were undergoing brain surgery under local anaesthesia. Libet took advantage of the discovery by Wilder Penfield and Herbert Jasper that conscious sensations of touch, referred to a particular place on the skin, can be evoked by electrically stimulating the surface of the somatosensory cortex.

Libet's initial discovery was that the train of stimulation pulses had to last for 200–500 milliseconds before the sensation of touch was perceived consciously. This delay — the 'mind time' — led him to propose his 'time-on theory', which states that conscious sensory experience requires that the appropriate brain activities turn on for at least 500 ms, and that shorter brain activities may lead to an unconscious mental function. This latter hypothesis is needed to explain fast reaction times, such as those of champion 100-metre sprinters who leave their blocks a mere 130 ms



Quick off the mark: Michael Johnson can be up and running before he is aware that the gun has fired.

after the starting pistol fires, suggesting that they are only aware that the gun has fired after they have left their blocks. But can the brain states of a patient undergoing brain surgery really be extrapolated to an athlete poised to break a world record?

In the visual system, electrical stimulation of the cortex in humans does not produce much more than tiny spots of light called phosphenes, as Giles Brindley discovered in the early 1960s. Bill Newsome and his colleagues found that direct electrical stimulation of the visual area called MT in monkeys influences the motion perception of the monkey only if it is combined with a real visual stimulus. The direct electrical stimulation on its own does not evoke a perception, or at least one that the monkey can report, so it has been difficult to use animal models to explore such issues.

Single impulses evoked in a sensory nerve by touching the skin can be detected, so one criticism of Libet's result is that when he stimulates the brain it evokes a rather coarse spatiotemporal pattern of cortical activation. The brain may simply be doing its best to make sense, literally, of the artificially induced activity of neurons that are normally activated from a patch of skin. This may take longer than usual, as the neural programs need to do the interpreting.

Libet's most controversial claim is that, because subjectively we believe that we are

immediately aware of the stimulus, there must be a disjunction between subjective timing and neuronal timing. On the basis of various experiments, he suggests that our sense of immediacy is an illusion that arises because we refer the timing of the experience back to the time of the earliest cortical activity, evoked some 10–30 ms after the skin is touched. The neural mechanisms by which all this precise time calculation has to be carried out is, of course, unknown.

Those of us for whom a slip of the tongue is a regrettably familiar event might be comforted to read here that, as in the case of sensory events, our conscious intention to act lags 300–500 ms behind the relevant brain events that actually lead to action. This of course raises the thorny question of who is actually doing the intending, as our conscious selves seem to be lounging in the back seat being chauffeured around by our zombie *doppelgänger*. The best we can do, it seems, is occasionally lunge for the steering wheel when our limousine takes a wrong turn. This is perhaps not quite as strange as it sounds, because highly practised movements, such as those of skilled musicians, are produced at rates that seem faster than thought.

In Libet's work, philosophers have found grist for what they do best. Indeed, his experiments, along with Larry Weiskrantz's discovery of blindsight, must rank as one

D. MILLS/AP PHOTO

of the major contributions of experimental psychology to modern philosophy of mind. Humans and monkeys with extensive damage to their primary visual cortex are clinically blind, but Weiskrantz found that they were surprisingly accurate about identifying the position and shape of objects when he forced them to guess. Philosophers, however, have not always been as appreciative of Libet's gift as they have of Weiskrantz's. Paul Churchland, Ted Honderich and Daniel Dennett, in particular, have attacked Libet's methods, results and ideas (and, in fairness, also each others'). What so vexes them is Libet's relegation of consciousness to the role of observer and occasional vetoer of action, which compromises both the cartesian view, in which the mind controls the machinery of the body, and the materialist view, where mind and brain events are identical.

Libet has replies for his critics, of course, but in developing his alternative theory, he ends up putting the ghost back in the machine in the form of a 'conscious mental field' emanating from the brain. This may be going too far, but whether or not one agrees with his thesis or not, one must acknowledge that his pioneering experimental work has certainly been stimulating. ■

Kevan Martin is at the Institute of Neuroinformatics, ETH/University of Zurich, Winterthurerstrasse 190, Zurich 8057, Switzerland.

A global problem

Global Change and the Earth System: A Planet under Pressure

by W. Steffen, A. Sanderson, P. D. Tyson, J. Jäger, P. A. Matson, B. Moore III, F. Oldfield, K. Richardson, H. J. Schellnhuber, B. L. Turner II, R. J. Wasson
Springer: 2004. 336 pp. \$129, £77, €99.95

Hans von Storch

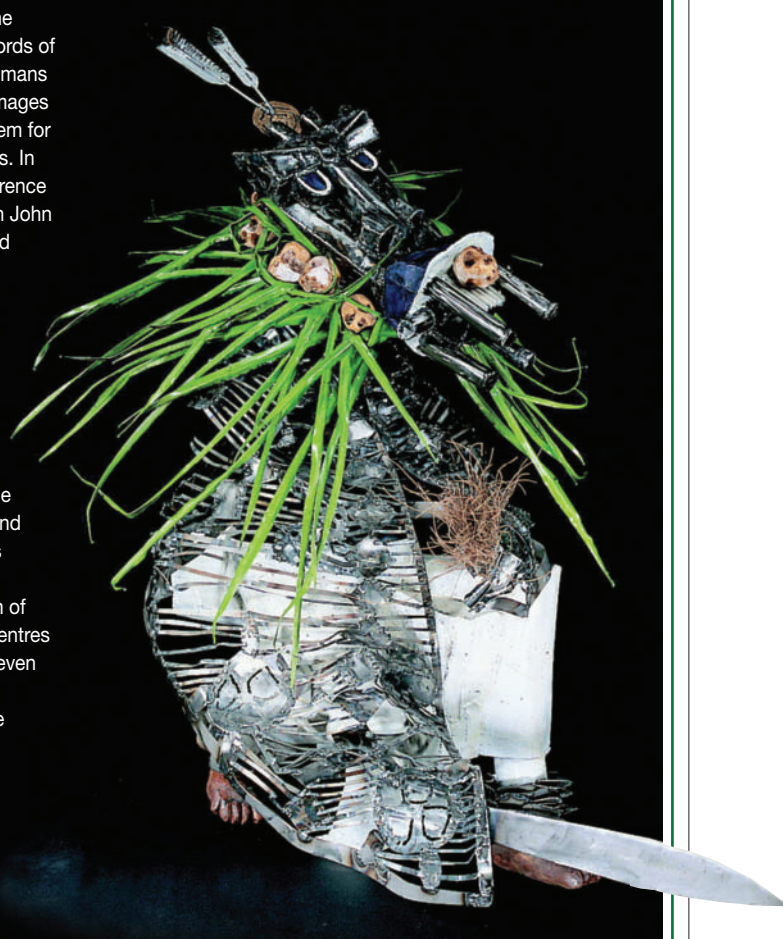
This book "focuses on the profound transformations of Earth's environment that is now apparent, a transformation owing not to the great forces of nature or to extra-terrestrial sources, but to the number and activities of people — the phenomenon of global change".

The authors have adopted a systematic approach. They begin by describing the Earth system — the whole of the physical and ecological environment that determines the state of the Earth — which they view as one integrated system. Next they discuss the dynamics of the Earth system as it was before human influence: the characteristics of temporal and spatial variability, the role of biological processes, the connectivity between the different parts of the system, and the role of nonlinearities, such as the browning of the Sahara. The third chapter describes the

A world of paint

Visual art has one of the longest and fullest records of any human activity. Humans have been recording images of the world around them for more than 30,000 years. In *Atlas of World Art* (Laurence King, £75), art historian John Onians has coordinated the work of a group of anthropologists, archaeologists and art historians to compile a temporal and geographic history of art.

Pictures of art, like the African masquerade shown here, take second place to detailed maps showing the cultural influences and location of key sites and artistic centres on each continent in seven different periods, from 40,000–5,000 BC to the twentieth century.



Anthropocene, the most recent part of Earth history, which has been, and continues to be, massively influenced by humans. Here the authors identify the drivers of change, characterize these changes, and put them into perspective in relation to the overall change of the Earth system. The fourth chapter then deals with the responses of the Earth system to human activities. The authors cover the entire spectrum of anthropogenic change, from emissions, changes in land use and land cover, and fires in South-east Asia, to the effect on coral reefs and the overall response of the Earth system with respect to climate and the water cycle, for example. They then examine the consequences of these changes for human well-being. The key points here relate to anticipation of future change, risks to key resources, and dangers to the Earth system as a whole. There is also a rather brief discussion of human perceptions of global change. In the final chapter, the authors draw various conclusions about the contemporary knowledge base, Earth-system science and the need for a toolkit with which to pursue the overarching goal of global sustainability.

The book is part of a series published for the International Geosphere-Biosphere

Programme (IGBP), one of four international global environmental-change research programmes sponsored by the International Council for Science. The IGBP's objectives are "to describe and understand the interactive physical, chemical and biological processes that regulate the total Earth System, the unique environment that it provides for life, the changes that are occurring in this system, and the manner in which they are influenced by human actions".

Global Change and the Earth System is richly illustrated with many useful coloured diagrams, and numerous boxes by additional authors provide more detailed information about specific issues. However, the wealth of detail makes it hard to read, and at times the diagrams are not properly explained.

Is it a good book? I am not entirely convinced. It deals with a broader set of topics (including the number of McDonalds restaurants around the world, and the emission of heavy metals) than the Intergovernmental Panel on Climate Change (IPCC) Assessment Reports. But the IPCC reports are written by a broader community of scientists, with more diverse expertise, greater emphasis on the certainty as well as the uncertainty of knowledge, and with more convincing

attempts either to explicate or fend off tacit value-driven concerns and opinions.

In *Global Change and the Earth System*, knowledge claims — that contemporary views are actually true — are often presented as definite explanations; examples include the huge extinction rate in Figure 1.7, and the reliance on the often-used but nevertheless questionable ‘wiggle matching’ technique. Another example is the claim that climate (in terms of an unexplained “Atmospheric Circulation Index”) is the only factor that affects catches of Japanese or European sardines, with social factors, such as the Second World War, playing only a minor role. The authors often adopt the unsatisfactory practice of relating a host of changes to ongoing anthropogenic change, without systematic attempts at formal detection and attribution.

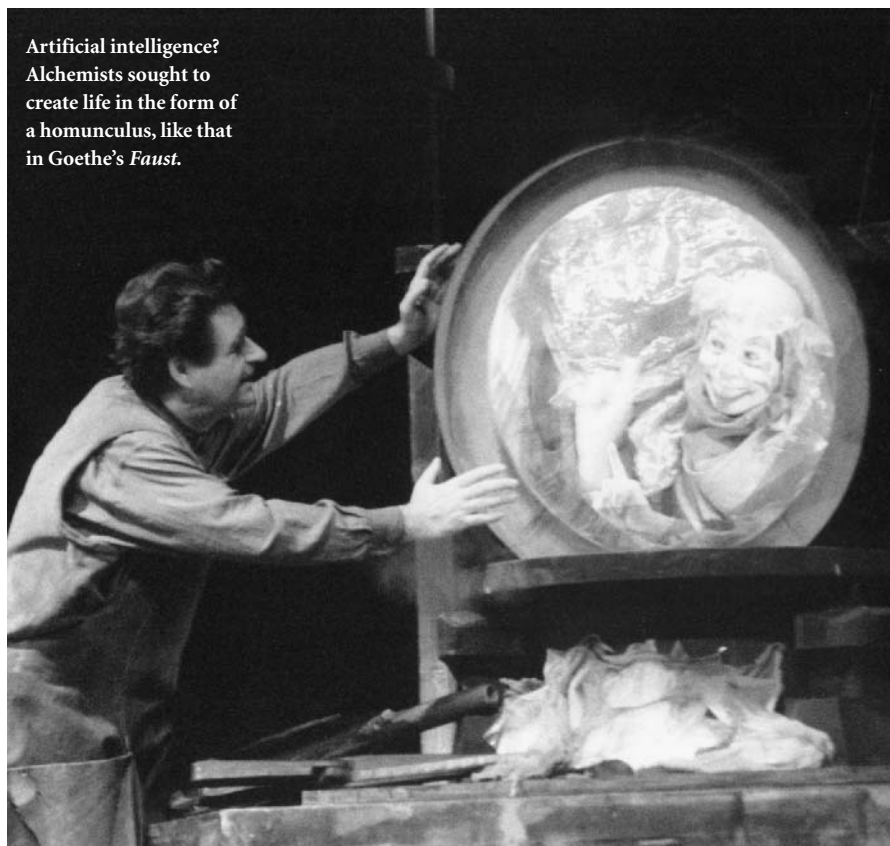
The examples that are adduced in this way seem to be biased towards deleterious effects. Thus, the book is a good demonstration that the environmental sciences are driven not only by curiosity and reductionist interests for detailed processes, but by an endeavour to integrate different sorts of often uncertain and contested knowledge claims from a broad field of disciplines — in an area that has high stakes. The result is an effort with a normative agenda of improving or saving the world.

I am sympathetic to this message, but as a scientist I think we must strive for objectivity (as far as possible) and avoid any overselling. Downplaying uncertainty is not useful, except perhaps for a limited time and a small audience. In the long term, this is not a sustainable approach. We can already see that large parts of the public and politicians, in both Europe and the United States, no longer trust many of the knowledge claims advanced by environmental scientists. Public concern about the state of the environment has begun to decline in recent years. Some may even erroneously see Bjørn Lomborg’s *The Skeptical Environmentalist* (reviewed in *Nature* 414, 149; 2001) as the other side of the coin to this IGBP volume.

The authors of *Global Change and the Earth System* believe that “any complete analysis of the consequences of global change must go well beyond scientific and economic considerations to fundamental moral and ethical values”. This is certainly true, but the book lacks an analysis of the complex history of the moral ideals in the cultural, ideological and political foundations of the modern environmental sciences. For this, readers can turn to such books as Ludwik Fleck’s *Genesis and Development of a Scientific Fact* (University of Chicago Press, 1979) or Clarence Glacken’s *Traces on the Rhodian Shore* (Cambridge University Press, 1967). ■

Hans von Storch is at the Institute for Coastal Research, GKSS Research Centre, Geesthacht 21502, Germany.

Artificial intelligence? Alchemists sought to create life in the form of a homunculus, like that in Goethe’s *Faust*.



Improving on nature

Promethean Ambitions: Alchemy and the Quest to Perfect Nature

by William R. Newman

University of Chicago Press: 2004. 352 pp.

\$30, £21

David Knight

There is nothing new under the sun, or so we are told. People worry about the artificial and the synthetic, as opposed to the natural; they pay extra for organic food, are concerned about genetic modification, and pontificate about bioethics, stem-cell research and cloning. All these things seem very modern, the fruit of progress in biochemistry, molecular biology and genetics; the oldest example would seem to be Justus Liebig’s advocacy of inorganic fertilizers in the hungry 1840s.

Not long before that, his friend Friedrich Wöhler had synthesized urea, in an attempt to provide evidence for atoms and their rearrangement, rather than to break down the organic/inorganic distinction. Processes and conditions in living organisms are very different from those in test-tubes, however, so although the end product might be the same, he did not replicate the reaction that creates urea in organisms. Human artfulness was still limited; perhaps nature was not after all a laboratory.

William Newman shows that debates about simulating, replicating and perfecting

nature go back much further than this, however. Whereas we come to these debates through modern chemistry and biology, medieval and early modern authors approached them through alchemy. Modern chemists tend to look at alchemy askance, as delusion, confidence trickery or occultism, leading to disgrace, beggary and early death. But some, Michael Faraday among them, had a degree of sympathy with the idea that all the metals could not be irreducibly different indestructible entities, and Ernest Rutherford cheerfully referred to his work as modern alchemy. Newman is prominent among the historians of science who have shown how important alchemy was as part of the serious ‘chymistry’ of Robert Boyle, Isaac Newton and their contemporaries. In this book, he looks at the divide between ‘art’, which used to mean anything productive involving artifice and forethought, and ‘nature’, as illuminated in discussions of, and laboratory and clinical practice in, alchemy.

If something were to be made from base metal that had all the properties of gold, would it be gold? Or is there some essence that distinguishes the natural from the ersatz? Are ‘species’, whether of metals or of creatures, God-given and immutable?

Astronomy, contemplating the starry heavens, was well-suited to natural theology, for God could be praised for His wisdom and grandeur. But alchemy, and the practical spin-offs that led to pigments, drugs and distillates, including alcohol, seemed to be an attempt to improve the world; God

Magnificent mayflies

Gaylord Schanilec is hooked on fishing flies that imitate nature.

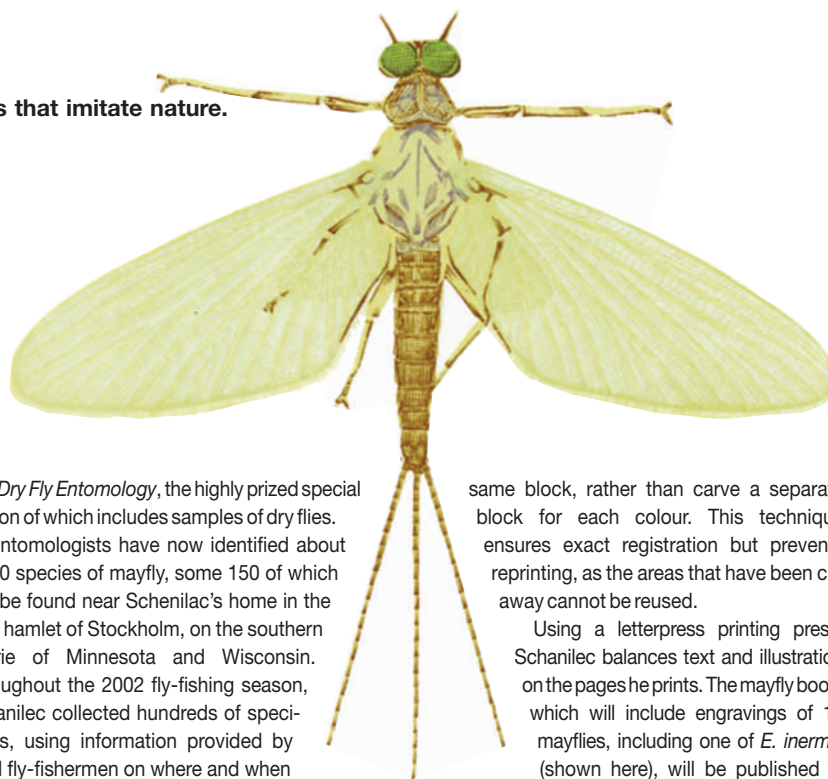
Colin Martin

Inspiration can be found anywhere. Gaylord Schanilec, wood engraver, fine printer and trout fly-fisherman found it on a river bank, in the form of the common mayfly.

Schanilec owns Midnight Paper Sales, a private press that publishes small, limited editions of illustrated books, often related to local history or geography. He designs, engraves and prints the books, and often writes the text. Each book project begins with an idea that evolves gently into the shape of the final book, a process that can take years. He is now hard at work illustrating and writing an as-yet untitled book on the species of mayflies (Ephemeroptera) found in his local rivers.

The idea for the book came to him in the summer of 2001 when, as usual, he spent as much time as he could fishing. This gave him plenty of time to contemplate the artificial fishing flies, hand-tied to imitate mayflies, with which he baited his hook. Trout are a major natural predator of mayflies in the wild, and trout fishermen have for centuries used local knowledge to tie appropriate artificial flies, called dry flies, as bait. This year, for example, is the sesquicentennial of Greenwell's Glory, a dry fly first tied by William Greenwell in 1854.

Not wishing to rely solely on fishermen's folklore, in the winter of 2001 Schanilec consulted Clarke Garry, a biologist at the University of Wisconsin at River Falls, who studies insect life in local trout streams. Under Garry's guidance, Schanilec bought a microscope, specimen vials and preservative agents, and developed techniques for collecting, preserving and documenting mayflies. He also read widely around his subject, including two late-nineteenth-century books by Frederick M. Halford, *Dry Fly Fishing in Theory and Practice*



and *Dry Fly Entomology*, the highly prized special edition of which includes samples of dry flies.

Entomologists have now identified about 2,000 species of mayfly, some 150 of which can be found near Schanilec's home in the river hamlet of Stockholm, on the southern prairie of Minnesota and Wisconsin. Throughout the 2002 fly-fishing season, Schanilec collected hundreds of specimens, using information provided by local fly-fishermen on where and when to find particular species, such as *Hexagenia limbata* on the lower Rush River and *Ephemerella subvaria* on particular stretches of the south branch of the Whitewater River. He has so far selected seven of these species for wood-cutting, including *E. subvaria*. Garry described them taxonomically and identified them for him.

Schanilec makes painstakingly detailed drawings. He then carves his final images into hard blocks of end-grain wood using sharp, metal gouging tools, to produce precisely detailed engravings from which he prints distinctive images in several colours. Reduction cutting — carving away more of the surface of the wood block between printings — allows him to print two or more colours from the

same block, rather than carve a separate block for each colour. This technique ensures exact registration but prevents reprinting, as the areas that have been cut away cannot be reused.

Using a letterpress printing press, Schanilec balances text and illustration on the pages he prints. The mayfly book, which will include engravings of 12 mayflies, including one of *E. inermis* (shown here), will be published in autumn next year in three editions: a large edition of portfolios of loose prints; a standard book, including both Garry's taxonomic descriptions and trout fishermen's folklore for each species bound alongside the prints; and a special edition, issued with hand-tied flies imitating each species, and an extra set of prints.

There is something rather magnificent about Schanilec's enterprise in capturing this most ephemeral order of insects between the covers of finely bound books. It recalls the remarkable eighteenth-century collaborations between artists and scientists in the Enlightenment.

Colin Martin is a writer based at 32 Woodstock Road, London W4 1UF, UK.

had left it to us to discover oil paints, milk of magnesia and whisky. We were (perhaps) better off as a result, but the natural theologian might be uneasy at the busy chymists' activities. And when naturally occurring substances were apparently made artificially, some were sceptical about whether they were exactly the same. Alchemists claimed that the processes by which gold was supposed to be generated in the bowels of the Earth were merely speeded up in their furnaces, but critics were not convinced, some seeing devilment lying behind any apparently successful transmutation.

Alchemists were not simply fixated on gold; they also sought not only to prolong life, but to synthesize it. Newman explores two routes to this end. The first was based on Jewish stories about the golem, made like Adam from the dust and brought to life (and destroyed when necessary) through Hebrew

words. A golem lacked a rational soul, and so was not human. The other tradition was a kind of male parthenogenesis, in which women (the source of flesh and blood) could be bypassed and a homunculus grown in a glass vessel from semen. The homunculus often appeared in alchemical pictures, and most famously in Johann Wolfgang von Goethe's play *Faust*. The homunculus was a spirit, unable to survive outside his flask; but from within it he could expound wisdom that was unavailable to ordinary mortals weighed down by their bodies. In this connection, Newman provides an interesting section on Paracelsus, the genius or charlatan who brought chymistry into medicine, with his sexual hang-ups and fascination with phoenix-like 'rebirths' through fire.

The ethical problems facing Newman's cast of characters were hypothetical — after all, alchemy didn't work. But in an age of test-

tube babies, Dolly the sheep and a vast array of synthetic chemicals, they are real enough.

Newman ends with a learned discussion of experiment — emphasizing that alchemy was an investigative science, carried on in the laboratory rather than the armchair — from medieval times up to Francis Bacon, Boyle, Newton and Margaret Cavendish. We meet artisans such as Bernard Palissy (who was 'making' fossils), as well as scholastic logicians. We are reminded that the 'scientific revolution' was not simply a matter of a new astronomy and a new physics. Rather, it was heavily dependent on the experimental tradition and critical thinking about analysis and synthesis. Alchemy would ultimately give birth not to a homunculus, but to the new chemistry. ■

David Knight is in the Department of Philosophy, University of Durham, 50 Old Elvet, Durham DH1 3HN, UK.

Last hideout of the unknown?

Scale and proportion: do the mechanisms of planar polarity also help determine the shape and size of animals?

Peter A. Lawrence

The mystery surrounding animal development is diminishing. We now understand much of how the DNA in the egg makes genes that build animals. We know how the body plan is elaborated from simple beginnings, and how the different types of cells are allocated and deployed. But there is one feature of animals for which we still have no explanation of any kind — yet it is so common that we pay it little attention. This is the characteristic and finely determined shape of animals — for instance, the three-dimensional space filled by a two-month-old okapi. A more familiar example is the fine sculpting of a face, such as the profile of your best friend. The mirror symmetry of the body, the trouble we have distinguishing identical twins and the way children resemble their parents should all remind us that these dimensions are prescribed genetically, written in the DNA sequence — but where and how? Here I conjecture, with only a little evidence, that the mechanisms that polarize cells within the plane of the epithelium could also help to sense dimension. An understanding of planar polarity might therefore help elucidate one of the deepest-remaining secrets of living things.

The important thing about growth is knowing when to stop. How do the cells in a growing bone of a mammal know when the final size has been reached? Moreover, both sides of an animal grow independently, but at identical rates — implying continuous action of the same control mechanisms on each side. In insects, the increase in size of the parts is spasmodic, being broken up into instars. At the start of the twentieth century, Harrison G. Dyar noted that the dimensions usually increase by the same proportion at each instar, giving a remarkably straight line on a log scale — again implying precise control. Some organs grow allometrically, that is, at a different rate to the rest. This precision suggests that there is a feedback from some measure of length in each main axis to every cell — for the decision to divide, not to divide or to die must be executed by each cell. The summed output of all those single-celled decisions in an organ will, in two or even three dimensions, determine its shape and size throughout growth and when it stops growing.

Any comparison between species raises the question of what is responsible for their different anatomies. I would guess that the main cause of evolution of shape is not the



Family resemblance: the shape and dimensions of our faces are prescribed in our genome.

changing of protein sequences one into another. For example, it does not occur simply because proteins such as haemoglobin differ slightly in sequence between man and mouse. Indeed, I suspect that if one were to take human proteins, one by one, and swap them with the homologous proteins in a mouse, most would do their job well. As increased numbers of proteins were substituted, I do not think that the result would be an increasingly human-looking mouse. This would mean that the anatomical differences between a human and a mouse would have to be sought elsewhere in the genome, in whatever control sequences are responsible for the length of the tail or the squeakiness of the larynx. Poodles and rottweilers, as well as the diverse shapes and sizes of beetles, illustrate how freely dimensions can be adjusted — by human or natural selection.

Dimension in one axis could be controlled by a monotonic gradient. If the limits of the gradient were fixed, then, as the axis grows longer, the gradient would become less steep across every cell — this could be how overall dimension is conveyed to each cell. For example, in the *Drosophila* abdomen, there are at least two gradient systems. One is an archetypal morphogen, the Hedgehog protein, which spreads outwards from a source. Its concentration fixes the pattern of cell types and the gradient of cell affinity. The other, the less well understood

gradient system, specifies planar polarity and seems to depend on intercellular bridges made by cadherin molecules. This polarity gradient might not decay with distance and could be of constant declination.

I offer three arguments that the polarity gradient, and not the morphogen gradient, specifies dimension. First, morphogen gradients do not seem to operate consistently over a growing field — as the axis elongates, some cells can become too remote from the source to be patterned. Large fields, such as developing butterfly wings, solve this problem by intercalating secondary morphogen sources to fill in the gaps. In the case of the *Drosophila* abdomen, regions far from the source are polarized normally and reach the right size even when rendered blind to Hedgehog — arguing that in this case Hedgehog is not directly responsible for either polarity or dimension. The second argument is that clones of cells that

lack the polarity cadherins not only have defects in polarity, but also grow excessively, suggesting that these cadherins are a link between planar polarity and growth. The third argument is one of mechanism. A cell in an epithelial sheet may resemble an amoeba of *Dictyostelium* that is becoming polarized by a gradient of cyclic AMP. The amoeba compares the amount of cyclic AMP reaching the periphery of the cell and moves up the gradient of concentration. Similarly, in a sheet of cells, each one could monitor its plasma membrane to detect the gradient and to read the vector that specifies its polarity. Could the cell also measure the amount of the difference across itself? If so, could this comparison give it information about dimension, and perhaps tell it when to stop dividing? ■

Peter A. Lawrence is at the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK. e-mail: pal@mrc-lmb.cam.ac.uk

FURTHER READING

- Coen, E. *et al. Proc. Natl Acad. Sci. USA* **101**, 4728–4735 (2004).
- Conlon, I. & Raff, M. *Cell* **96**, 235–244 (1999).
- Day, S. J. & Lawrence, P. A. *Development* **127**, 2977–2987 (2000).
- Hafen, E. & Stocker, H. *PLoS Biol.* **1**, E86 (2003).
- Lawrence, P. A. *The Making of a Fly: the Genetics of Animal Design* (Blackwell, Oxford, 1992).
- Thompson, D. W. *Growth and Form* (Cambridge Univ. Press, Cambridge, 1942).

Pillow talk in plants

Bruce McClure

Flowering plants have sophisticated mating systems that determine whether a pollen suitor is accepted or rejected. Knowledge of how such systems operate has just taken two steps forward.

Few *Nature* readers may admit to such a lack of subtlety, but most will have closely scrutinized potential mates. Mate selection has a huge effect on a species. In general, it's good to avoid inbreeding and it's bad to mate with other species. Humans possess sophisticated sensory faculties that help make these selections. Flowering plants have the same mating imperatives, but they lack our sensory faculties, are largely immobile, and rely on the vagaries of factors such as wind or insects to assist mating. Their remarkable success is due, in part, to the evolution of sophisticated mating control systems that require communication between pollen (the male player in events) and pistil (the female organs, consisting of a stigma on which the pollen lands, the egg-containing ovary, and a style to provide passage from one to the other).

Plants alternate between a diploid generation, with two copies of each gene, and a haploid generation, the gametophyte (the generation that produces gametes, the sperm or egg), with only one copy of each gene. In flowering plants the diploid generation is visible and familiar to us; the male and female gametophytes are not because they are microscopic. But what mostly sets flowering plants apart from other plants is that their eggs are enclosed in the diploid tissues of the pistil. Pollen must germinate and produce a pollen tube to carry the haploid sperm cells through these tissues. This requirement provides an opportunity for a two-sided conversation in which a decision can be made about whether pollen should be accepted or rejected. Two papers^{1,2} in this issue report major insights into how the pollen–pistil communication systems work.

Studies of self-incompatibility mechanisms have been the most effective way of listening in on these conversations. Self-incompatibility allows plants to avoid inbreeding by rejecting pollen from close relatives and, in some cases, it contributes to rejecting pollen from other species as well. The result is that genetic diversity and fitness are maintained. Compatibility is genetically controlled by an 'S-locus' carrying distinct specificity genes, one expressed in the pollen and the other in the pistil. The S-locus is extremely polymorphic; even small populations may have dozens of different S-haplotypes (S_1 , S_2 , and so on). Gametophytic self-incompatibility is the most common



Figure 1 Choosy mates. This field poppy and other self-incompatible plants can distinguish between cross-pollen from an unrelated plant and self-pollen that has travelled only from the ring of yellow anthers to the central pistil. Pollen–pistil communication allows exquisite control over plant mating.

type of genetic control. Haploid pollen is rejected when its S-haplotype is the same as either of the two S-haplotypes in the diploid pistil. Any other combination is compatible. The challenges are to identify the gene products that bring about S-specific pollen rejection and to determine how they function.

In one of the new papers (page 305), Thomas and Franklin-Tong¹ address function: they show that gametophytic self-incompatibility in the field poppy (*Papaver rhoeas*; Fig. 1) involves programmed cell death. In poppies, the conversation opens on the pistil side with S-proteins produced on the stigma surface. Pollen that does not share an S-haplotype with the pistil is compatible. The conversation goes well for such pollen, and access to the ovary is permitted. For incompatible pollen (that is, an S-haplotype that matches the pistil), the conversation goes poorly. The S-proteins interact with an as-yet-unidentified pollen factor, resulting in immediate inhibition of pollen-tube growth and — as Thomas and Franklin-Tong now show — initiation of events

that are typical of the programmed-cell-death response (release of mitochondrial cytochrome c, generation of a caspase-like enzyme activity and DNA fragmentation). These results reveal a whole new landscape of events on the pollen side, and show just how badly such a conversation can end for an undesirable mating partner.

Although the genetics of self-incompatibility in the potato, rose and foxglove families (Solanaceae, Rosaceae and Scrophulariaceae, respectively) are the same as in the poppy, the biochemical details of the conversation are completely different. It is known that self-incompatible plants in these families have a system that involves an S-specific inhibition of incompatible pollen-tube growth, one that involves a cytotoxic — cell-killing — response. But our knowledge of the system has been one-sided³. Each different S-haplotype carries a unique *S-RNase* gene, which encodes an enzyme capable of degrading RNA and inducing a cytotoxic response. S-RNases are secreted along the path where the pollen tube grows through the pistil.

NONI FRANKLIN-TONG

They are thought to act as S-specific cytotoxins because their enzyme activity is required for pollen rejection and because pollen RNA becomes degraded in incompatible pollinations. The genetics of gametophytic self-incompatibility require that both the pollen and the pistil contribute to the conversation. Until now, the pollen S-gene, the pollen specificity gene, has been elusive. In the other new paper (page 302), Sijacic *et al.*² describe how they have found it.

Several groups have identified pollen-expressed genes that meet the genetic requirements for a pollen S-gene candidate — linkage to the S-locus, expression in pollen and appropriate polymorphism. Like the other groups, Sijacic *et al.* sequenced the genomic region around the *S-RNase* gene and found a pollen-expressed, haplotype-specific gene close to the *Petunia inflata* *S₂-RNase* gene. But they took their studies a step further and directly tested whether their candidate pollen S-gene (*PiSLF₂*, for *Petunia inflata* S-locus F-box *S₂*-haplotype) determines haplotype specificity in pollen. This turns out to be a subtle matter. The only option was to test for an interaction between *PiSLF₂* and pollen S-genes in different haplotypes. *PiSLF₂* behaved exactly as predicted, and the results conclusively show that *SLF* is the long-sought pollen S-gene.

The pollen S-gene sequence provides clues to its function in self-incompatibility. It has long been thought that pollen S functions as an inhibitor of S-RNase in compatible pollinations. Several lines of evidence support this idea, and it is compelling because of the pressure on the pollen to overcome fertilization barriers in the pistil. The SLF protein contains a motif, called the F-box, which is best known for mediating interactions with other proteins that make up an enzyme complex, the E3 ubiquitin ligase complex. Ubiquitination of cellular proteins serves a variety of functions, including acting as a degradation or targeting signal. The cytotoxic activity of S-RNase requires an active enzyme with access to the cytoplasm. Therefore, SLF could provide resistance to S-RNase either by stimulating degradation of S-RNase or by preventing its access to the cytoplasm.

Qiao *et al.*⁴ have shown that *Antirrhinum* *AhSLF₂* binds to S-RNase and interacts with components of the ubiquitination complex, and they suggest that S-RNase is degraded in compatible pollinations. However, there is abundant S-RNase in both compatible and incompatible pollen⁵, so it is not yet certain that degradation of S-RNase is the basis for resistance in compatible pollen.

Nonetheless, with the identification of pollen S we can finally listen in on both sides of the pollen–pistil conversation. The pistil secretes cytotoxic S-RNases to inhibit undesirable pollen and SLF provides resistance to that except when the pollen and pistil have

matching S-haplotypes. Given the new findings, research can now shift towards decoding the biochemical details of the exchanges between pollen and pistil. ■

Bruce McClure is in the Department of Biochemistry, University of Missouri-Columbia, Columbia, Missouri 65211, USA.

e-mail: mcclureb@missouri.edu

1. Thomas, S. G. & Franklin-Tong, V. E. *Nature* **429**, 305–309 (2004).
2. Sijacic, P. *et al.* *Nature* **429**, 302–305 (2004).
3. Kao, T.-h. & Tsukamoto, T. *Plant Cell* advance online publication 9 March 2004 (doi:10.1105/tpc.016154).
4. Qiao, H. *et al.* *Plant Cell* **16**, 582–595 (2004).
5. Luu, D. T. *et al.* *Nature* **407**, 649–652 (2000).

Planetary science

A perfect match?

Beth Ellen Clark

The 'S-complex' asteroids are not easily identified as the source of the most common meteorites reaching Earth. Their relationship might be disguised, however, by the effects of space weathering.

Meteorites that fall to Earth come primarily from asteroids — those minor planets that orbit the Sun mainly between Mars and Jupiter. We have meteorites from approximately 100 different asteroid parent bodies, and the great challenge in meteorite–asteroid studies is to determine which meteorites come from which asteroids. A basic correlation has been established between types of meteorites and types of asteroids based on their 'colour' — the amount of light reflected at one wavelength compared with the amount reflected at another. Telescopic observations of asteroid colours and laboratory measurements of meteorite colours suggest relations between them, and have also revealed much about the environment of the early Solar System in which the asteroids formed.

It is therefore terribly perplexing that there is abundant other evidence supporting a link between the most common meteorite type, the ordinary chondrites, and the most common asteroid type in the inner main belt, the S-complex asteroids — yet their colours don't match! One suggested solution to the problem is that the colour of these asteroids may have changed as their surfaces became weathered in the space environment. Other solutions have been proposed, but they create more problems than they solve; colour change would be Occam's choice. Now Jedicke *et al.*¹ demonstrate an interesting correlation between the age of an asteroid's surface and its colour (page 275 of this issue), and use it to argue in support of the space-weathering process.

It is clear that, over time, the exposure of airless bodies to the environment of space will change the optical properties of their surfaces. Most pertinent evidence comes from studies of lunar soils^{2,3}, backed up by results from spacecraft observations⁴ and from experiments with meteorites⁵. The surfaces of asteroids are subject to impacts, ion-implantation from the solar wind, sputtering and micrometeorite bombardment — effects that collectively have come to be

known as space weathering (Fig. 1). But whether this is the explanation for the colour mismatch between the ordinary chondrites and the S-complex asteroids is controversial.

Jedicke *et al.*¹ have found a way to address the conundrum, based on the age of asteroid families. The orbital characteristics of asteroids within a family are so similar that the family members are inferred to be pieces of a once-single object that broke apart through collisions. This disruption would presumably have brought fresh bedrock material to the surface of each fragment. Then the asteroid family would gradually change colour through space weathering, starting from the time of its disruption. Jedicke *et al.* have shown that the age of an asteroid family is directly correlated with average colour in exactly the sense predicted by models of space weathering.

What else could explain this correlation? It is possible that the dating technique is erroneous, or that the strong correlation seen for 11 families (out of 12 studied) arises because their mineralogical composition is such that their colours just happen to vary in proportion to their age. But this is grasping at straws. Nevertheless, it is somewhat troubling that there is no significant concomitant correlation of albedo with asteroid-family age. Albedo is the fraction of light reflected by the object, and space-weathering effects are clearly associated with a decrease in albedo.

In any case, Jedicke *et al.*¹ present an interesting interpretation of the correlation, and estimate the rate of space weathering for asteroids. This rate turns out to be about 100 times longer than any other estimate on the books. Furthermore, when they extrapolate back to a time shortly after the formation of each asteroid family, the predicted colour of the asteroid parent body matches that of ordinary chondrites: the implication is that these meteorites are chips off the blocks of S-complex asteroids. Certainly, this work will cause a stir in the trenches of asteroid science.

Although other meteorite–asteroid links

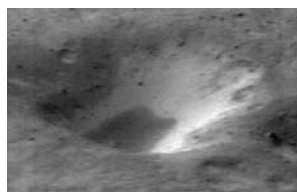


Figure 1 Weather report. Left, asteroids such as 433 Eros, captured here by NASA's Near Earth Asteroid Rendezvous (NEAR) Shoemaker mission, show evidence of space weathering. Above, close-up of a crater on Eros, showing bright subsurface material that has been exposed by downslope movement of space-weathered surface material.

are affected by colour changes, the link between ordinary chondrites and S-complex asteroids is the single most important example of the problem because of their great abundance: ordinary chondrites account for more than 80% of meteorites; S-complex asteroids make up more than 80% of some parts of the main belt.

A definite link between the ordinary chondrites and the S-complex asteroids would have important implications for meteorite and asteroid science, and for studies of the early Solar System. If these asteroids are made of ordinary-chondrite material, then such material must have been very common indeed and might even have been the starting material from which the terrestrial planets formed. If S-complex asteroids are not made of ordinary-chondrite material,

then dynamicists have a lot to do to explain how the process that brings meteorites from the asteroid belt to Earth could have overwhelmingly over-sampled a rare type of asteroid.

I suspect that, until a spacecraft can scoop up a surface sample from an S-type asteroid and return it to Earth, the debate will continue, with each new piece of evidence adding fuel to the fire.

Beth Ellen Clark is in the Department of Physics, Ithaca College, Ithaca, New York 14850, USA. e-mail: bclark@ithaca.edu

1. Jedicke, R., Nesvorný, D., Whiteley, R., Ivezić, Ž. & Jurić, M. *Nature* **429**, 275–277 (2004).
2. Hapke, B. J. *Geophys. Res.* **106**, No. E5, 10039–10073 (2001).
3. Pieters, C. M. et al. *Meteor. Planet. Sci.* **35**, 1101–1107 (2000).
4. Clark, B. E. et al. *Meteor. Planet. Sci.* **36**, 1617–1638 (2001).
5. Sasaki, S., Nakamura, K., Hamabe, Y., Kurahashi, E. & Hiroi, T. *Nature* **410**, 555–557 (2001).

Condensed-matter physics

Atomic beads on strings of light

Murray J. Holland

A new regime of strongly correlated quantum behaviour has been reached with the creation of a one-dimensional Tonks–Girardeau gas from ultracold atoms trapped within thin tubes of light.

Some materials reveal their underlying quantum nature at ultracold temperatures. Spectacular examples are superconductors and superfluids, in which frictionless currents flow at temperatures near absolute zero. It is quantum mechanics that brings about these effects, through particles condensing into a 'superfluid' quantum state. The resulting properties are strongly dependent on the dimensions in which the superfluid can move: in three dimensions, the motion is unrestricted; in two dimensions, the flow is confined to sheets or surfaces; and in one dimension, to lines or strings. On page 277 of this issue, Paredes *et al.*¹ report a remarkable experiment in which they created a one-dimensional quantum

gas, trapped within thin tubes of light. The interactions between the trapped ultracold rubidium atoms were enhanced by the confinement and could be modified, bringing about a continuous passage from a weakly interacting regime to strongly correlated behaviour in what is known as a Tonks–Girardeau gas^{2,3}.

In quantum mechanics, particles are classified into two types — bosons and fermions. The classification is based on what happens to the 'wave function', the quantum mechanical description, of two identical particles when the particles are exchanged, one for the other. For exchanged bosons — two rubidium atoms, for example — the wave function is unchanged, but for fermions the wave

function gains an overall minus sign. One consequence is that identical fermions can never be at the same point in space: swapping the particles would have no effect, but would still introduce a minus sign into the wave function; assuming the wave function is non-zero, there is then an inconsistency, as the same state would have two different wave functions.

Could bosons be made to behave in this fermionic way? The most extreme situation would be a row of impenetrable bosons on a one-dimensional string; they cannot occupy the same location and are therefore unable to pass each other and exchange places. The resulting traffic jam would involve strong correlations — the motion of each boson would be correlated with that of the next one in the line. Yet even in this case, their behaviour would not be exactly like that of fermions. The fermionic exchange rule implies more than the exclusion of two particles from the same point, as seen from the fact that the momentum of two identical fermions can never be the same, regardless of where they are located. However, in practice, so many properties of this one-dimensional string of bosons would be fermion-like that the situation is often referred to as the 'fermionization' of bosons.

That statement can even be made mathematically precise. More than forty years ago, it was realized that there is an exact one-to-one mapping of impenetrable bosons in a one-dimensional system onto a system of fermions that do not interact at all³. In fact, this problem connects with a general area of theoretical physics. Calculations showed that the lowest energy excitations of the system were similar to those of sound waves⁴. This in turn provided a powerful link to a universal class of 'Luttinger liquids'⁵, describing many one-dimensional models in which the excitations have this character.

So how could such an interesting system be realized? Until recently, the discussion has been largely academic because of the lack of appropriate physical systems. But developments in the field of quantum gases have provided a natural starting point. The production of ultracold superfluid gases is now routine, and low-dimensional systems can be created using optical lattices. An optical lattice is a web of intersecting laser beams, generating a three-dimensional interference pattern. Atoms inside this web are exposed to a frictionless field of electromagnetic potential, whose spatial structure follows the intensity of the light. Many lattice structures are possible, but the relevant one here is a matrix of extremely thin, parallel tubes, which confine the atoms to one-dimensional motion.

Constricted in this way, the interactions between the atoms are then determined by only a few parameters, including their collision properties in free space⁶, their atomic

Malaria

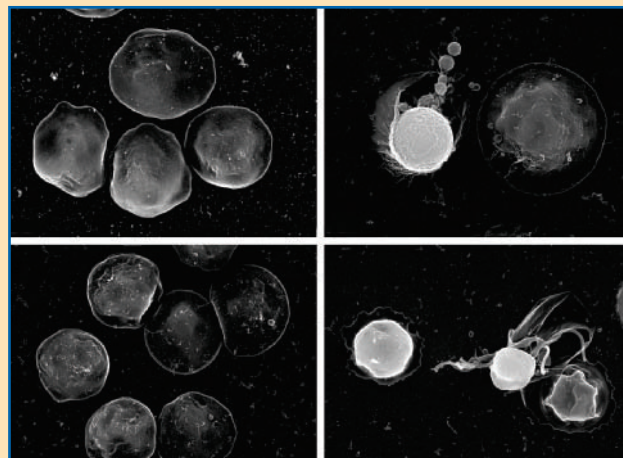
The calcium connection

One of the most visually striking events in the life cycle of the malaria parasite is exflagellation — the process by which the male reproductive cells emerge, complete with propellers, from a precursor cell. As they describe in *Cell* (117, 503–514; 2004), Oliver Billker and colleagues have pinned down the initial steps in the underlying cascade of molecular events.

The malaria parasite leads a complex life, relying on two hosts: humans and mosquitoes. It reproduces sexually in the insects. When a mosquito bites an infected human, it takes up red blood cells containing parasite gametocytes — pre-reproductive cells. Within seconds of being ingested, the gametocytes break out of their unwitting host cells, destroying them in the process. The resulting female reproductive cells (gametes) can be fertilized immediately, but the male cells must go through several more stages of development before fully fledged male gametes are formed.

Like all good parasites, the malaria agent takes advantage of its host's own machinery to spark off these events — the mosquito molecule xanthurenic acid acts as the trigger. But what happens next? Billker *et al.* chose to investigate the role of Ca^{2+} ions, common molecular mediators in many organisms, and Ca^{2+} -binding proteins, of which there are many encoded in the parasite genome. To do so, the authors engineered *Plasmodium berghei* — a rodent malaria parasite — to carry a molecule that glows brightly in the presence of Ca^{2+} ions. When they added xanthurenic acid to mouse blood containing these engineered parasites, the gametocytes lit up (the Ca^{2+} detected probably coming from stores within the parasites).

So what occurs downstream of the release of Ca^{2+} ions in the production of gametes? Billker *et al.* wondered if a protein of the calcium-dependent protein kinase (CDPK) family might be implicated.



Indeed, the authors found that one member of the family, CDPK4, is expressed predominantly in male *P. berghei* gametocytes. They then went on to generate parasites lacking this protein. As the top two images show, mutant male gametocytes were able to emerge successfully from red blood cells (left, red blood cells; right, a rounded male gametocyte emerging from the ghost of its host). But they could not

exflagellate: to do so, CDPK4 had to be reinstated (see the lower two images; an exflagellating gametocyte, with protruding propellers, is on the right).

It remains to be seen what comes next in this game of molecular dominoes. Billker *et al.* also want to know where exactly the Ca^{2+} ions are stored inside gametocytes — and is there a counterpart of CDPK4 in the female cells?

Amanda Tromans

V. BRINKMANN/ELSEVIER

mass and the thickness of the light tubes. Perhaps surprisingly, the effective mass of each atom can be easily changed by introducing a standing-wave light field along the length of the tubes, giving rise to a periodic one-dimensional potential. The resulting motion, in which an atom continuously absorbs and emits light as it surfs the standing wave, modifies the effective atomic mass in a way that can be controlled. Using this technique, Paredes *et al.*¹ observed the cross-over⁷ from the weakly interacting regime, in which the particles can freely pass one another, to a 'fermionized' strongly correlated gas. As proof of the creation of the Tonks–Girardeau gas, Paredes *et al.* measured the momentum distribution of the atoms in the tubes and found that it matched their theoretical prediction for such a state.

It should be emphasized that the Tonks–Girardeau gas only emerges if the gas is sufficiently dilute. At high density, another strongly correlated state with quite different properties may be produced. This is the Mott insulator, which was recently studied in a similar but complementary experiment⁸.

One of the most important aspects of Paredes and colleagues' experiment is the striking illustration of quantum control and quantum-state engineering. How this will

affect the ongoing pursuit of quantum information and computation, and the investigation of a variety of quantum coherent and strongly correlated phenomena, is not yet clear. But an exciting path has been laid out before us.

Murray J. Holland is in the Department of Physics, University of Colorado, Boulder, Colorado 80309-0440, USA.

Gene regulation

Selfish elements make a mark

Frederic Bushman

Transposons qualify as 'selfish' DNA elements, adding new copies of themselves into our genomes without regard for the consequences. This wilful habit may, however, help in normal gene regulation.

People differ in their genetic make-up, not only through changes in single DNA base pairs but also by the insertion of whole blocks of DNA^{1–3}. Human cells contain one type of active mobile element — the L1 retrotransposon — that is responsible for these insertions. Transposons are defined as DNA sequences that can form extra copies of themselves at new locations in an

e-mail: murray.holland@colorado.edu

1. Paredes, B. *et al.* *Nature* **429**, 277–281 (2004).
2. Tonks, L. *Phys. Rev.* **50**, 955–963 (1936).
3. Girardeau, M. *J. Math. Phys.* **1**, 516–523 (1960).
4. Lieb, E. H. & Liniger, W. *Phys. Rev.* **130**, 1605–1616 (1963).
5. Haldane, F. D. M. *J. Phys. C* **14**, 2585–2609 (1981).
6. Olshanii, M. *Phys. Rev. Lett.* **81**, 938–941 (1998).
7. Petrov, D. S., Shlyapnikov, G. V. & Walraven, J. T. M. *Phys. Rev. Lett.* **85**, 3745–3749 (2000).
8. Stöferle, T., Moritz, H., Schori, C., Köhl, M. & Esslinger, T. *Phys. Rev. Lett.* **92**, 130403 (2004).

organism's genome; today, L1 and related retrotransposons account for a staggering 17% of the human genome sequence. Mercifully, however, L1 elements 'transpose' only infrequently, minimizing damage to our chromosomes. The forces that keep L1 at bay have long been a mystery, but two papers by Boeke and colleagues in this issue^{4,5}, together with an earlier study⁶, unveil a major control



100 YEARS AGO

A copy of the *Peterborough Advertiser* of May 7 has been sent to us, containing the announcement that radium has been found in beds of Oxford Clay near Fletton, Huntingdonshire. No particulars are given, but a long descriptive article on the discovery suggests that it will make "brickfields better than gold mines." These sanguine anticipations will perhaps be tempered by the following extract from a paper by Prof. J. J. Thomson, read before the Cambridge Philosophical Society on February 15:—"Radium was found in garden soil from the laboratory garden, in the Cambridge gault, in gravel from a pit at Chesterton, in still greater quantities in sand from the sea-shore at Whitby, in the blue lias at Whitby, in powdered glass, in one specimen of flour, and in a specimen of precipitated silica."

ALSO

In the souring of milk the amount of lactic acid developed may reach 0.80 per cent. in three or four days when the milk solidifies. In view of Sir O. Lodge's suggestion (*NATURE*, October 1, 1903), I have made experiments comparing the rate of acidification, in two to three days, with and without the influence of radium rays... It therefore appears to me that under normal conditions radium rays have little or no effect on the functions of the lactic acid bacillus.

From *Nature* 19 May 1904.

50 YEARS AGO

Mr. A. R. Thomson, R. A., has completed a painting on astronomy to decorate a wall in the entrance hall of the Science Museum, London, near the well-known Foucault pendulum. Mr. Thomson was given *carte blanche* to depict the subject of 'Astronomy', and he has chosen as his main theme the invention and development of the telescope. On the left the small sons of Dutch spectacle-makers are playing with combinations of pairs of lenses... In the centre of the picture Galileo is using his telescope, ministered to by three rather bewildered ladies. On the right the distrust and suspicion of this early period is shown by figures of authority against a background of burning books. The whole picture is dominated by the great two-hundred-inch Hale telescope, and the sky behind is an animated representation of heavenly bodies among the figures of the constellations.

From *Nature* 22 May 1954.

mechanism. In the first paper⁴, the authors present this discovery and discuss its implications for understanding genomic evolution. Their second paper⁵ describes how they capitalized on these findings to make a modified L1 element that is more than 200 times as active as normal.

L1 and its relatives are termed retrotransposons because their transposition involves reverse transcription — the copying of RNA into DNA. Retrotransposition of L1 begins with transcription of the L1 genomic (DNA) sequence to form messenger RNA (mRNA) copies, a process carried out by a cellular enzyme called RNA polymerase (Fig. 1). The resulting RNAs serve both as templates for the production of L1 proteins (ORF1p and ORF2p) and as templates for retrotransposition. ORF1p binds and organizes the L1 RNAs, whereas ORF2p contains two enzymatic activities — the reverse transcriptase and an endonuclease.

The endonuclease makes a DNA nick in a target chromosome (Fig. 1). The reverse transcriptase then extends one of the freed DNA ends, using the L1 RNA sequences as a template, thereby making a single-stranded DNA copy of the L1 RNAs at the nicked site. Next, the new L1 DNA sequence becomes double-stranded and fully integrated through a series of, as yet ill-defined, further steps. As a result, a copy of the L1 sequence appears at a new location in the genome — often within a gene. The nicking reaction is weakly sequence-specific (it favours DNA that is rich in adenine and thymidine bases), but the target sequence is common enough for L1 elements to be able to hop into much of the human genome.

Nonetheless, the rate of L1 transposition today is very low. This is because the L1 RNAs and proteins accumulate only to very low levels in non-reproductive cells (the proteins are expressed more efficiently in germline cells, explaining why people often differ in the insertions in their genomes). In the first of the new papers, Han, Szak and Boeke⁴ describe how they tracked this weak accumulation down to barriers within the L1 DNA sequence that prevent the efficient passage of RNA polymerase. They fused the ORF1 or ORF2 DNA sequence to marker genes, and showed that the hybrid mRNA accumulated much less than did the marker gene mRNA alone.

Curiously, this effect was independent of the orientation of the L1 DNA — but the mechanism varied. When the L1 sequence was in the same orientation as the marker gene, transcription began but elongation of the mRNA was inhibited. In the opposite orientation, the polyadenine — poly(A) — sequence that is characteristic of the tail end of mRNAs was added prematurely (a mechanism that has been suggested previously⁶).

Boeke and colleagues⁴ also carried out a bioinformatics study of the human genome,

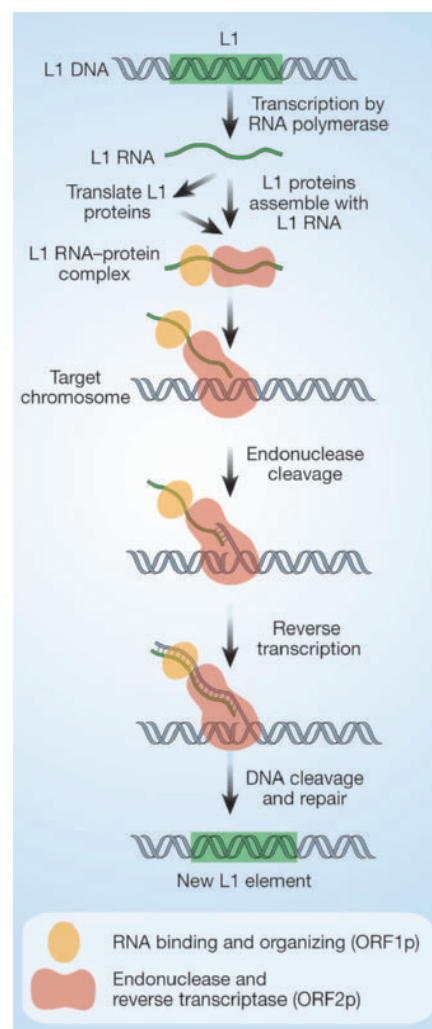


Figure 1 Retrotransposition of the L1 element. The process begins when the DNA sequence encoding an L1 element is transcribed into RNA by RNA polymerase; the L1 proteins ORF1p and ORF2p are then made, using the RNA as a template, and the RNA and the proteins form a complex. The complex binds to another part of the genome and the endonuclease makes a nick in the DNA, generating a free end. The reverse transcriptase uses this free end as a 'primer', allowing it to start synthesizing a single-stranded DNA copy of the L1 RNA. The RNA is then removed (not shown) and the single-stranded DNA is used as a template to make double-stranded DNA, which is then seamlessly stitched into the nick in the chromosome. Thus, a new L1 element has been made in the genome.

which revealed that these obstacles can be imposed on a cellular gene through the integration of L1 sequences. Specifically, the authors found that some 79% of human genes have L1 sequences in their non-protein-coding regions (introns) — and that high L1 density correlates with low levels of accumulation of the mRNAs encoded by these genes. These findings support a model in which L1 insertions in introns regulate cellular gene activity by diminishing the efficiency of mRNA elongation or promoting

the premature addition of poly(A) tails, usually without disrupting the integrity of the final protein.

Armed with these findings, Han and Boeke⁵ created a mouse hyperactive L1 element (smL1) by resynthesizing the coding sequence of L1. This striking technical feat involved assembling 4,500 base pairs of continuous DNA from short synthetic segments. To make smL1, they replaced a large fraction of the L1 element's protein-coding region with 'synonymous' sequences — sequences that encode the same amino acids but use different triplets of bases to do so. They thereby mutated the sequences that prevent transcriptional elongation or allow premature addition of poly(A) tails. Transposition tests in cultured cells revealed that smL1 transposed more than 200 times more efficiently than wild-type L1.

These findings have several implications. One is technical — if smL1 is as active in intact animals as it is in cultured cells, it might provide a new tool for generating mutations in genetic screens. Human L1 elements have been shown to transpose in mice⁷, but use of smL1 instead could greatly increase the rate of mutagenesis. This strategy has the advantage that mutations caused by inserting smL1 would be tagged with smL1 sequences, making it easier to isolate the mutated gene.

Another implication involves the long-term remodelling of the host genome. L1 elements qualify as 'selfish DNA', responding individually to darwinian selection. But the genetic novelty they generate can contribute positively to their host as well, by providing new substrates for natural selection⁸. Previously, researchers have shown that L1 elements can modify the genome by, for example, moving adjacent sequences about with them, hopping into (and so mutating) the coding portions of genes, and contributing portable regions of similarity that allow chromosomes to pair up and swap segments.

The new papers^{4,5}, together with the earlier study⁶, suggest that the insertion of L1 elements into introns can also diminish cellular gene expression in a graded fashion. In the words of Han, Szak and Boeke, such L1 insertions provide a "molecular rheostat" with which to govern gene activity — and their bioinformatics analysis establishes that the mechanism is widely used. Given the similarity of gene complements among species, it seems likely that relatively modest changes in gene activity are a key feature of many speciation events. The new findings may thus have revealed a major mechanism that connects genomic plasticity to the evolution of whole organisms. ■

Frederic Bushman is in the Department of Microbiology, University of Pennsylvania School of Medicine, 3610 Hamilton Walk, Philadelphia, Pennsylvania 19104-6076, USA.

e-mail: bushman@mail.med.upenn.edu

1. Boissinot, S., Chevret, P. & Furano, A. *Mol. Biol. Evol.* **17**, 915–928 (2000).
2. Martin, S. L. & Garfinkel, D. J. *Genome Biol.* **4**, 313 (2003).
3. Kazazian, H. H. *Science* **303**, 1626–1632 (2004).

Chemistry

Atom tracking

Stuart A. Rice

Diffraction patterns, generated using X-ray pulses of only picosecond duration, reveal the motion of individual atoms as iodine molecules dissociate, then recombine, in solution.

Location, location, location! That well-known saying about real-estate value is an apt paraphrase of what chemists want to know about chemical reactions, as atoms change position to transform one molecule into another. These changes govern the outcome of a reaction, and ultimately determine the biological function of the molecule. As they report in *Physical Review Letters*, Anton Plech *et al.* have extended the technique of X-ray diffraction to track such changes on a shorter timescale than ever before (*Phys. Rev. Lett.* **92**, 125505; 2004). They have monitored the photodissociation of iodine dissolved in carbon tetrachloride at the level of picoseconds — just 10^{-12} seconds.

The general outline of a scheme to track a chemical reaction has been clear for some time: generate a short-lived excitation of the molecule to initiate a reaction, and then, over varying time intervals, follow the changing atomic structure of the reactant. Although suitable excitation sources — such as lasers that generate femtosecond (10^{-15} s) pulses of tunable light — have been available for more than a decade, only recently has it become possible to record the change in structure of a reacting molecule with appropriate time resolution. The development of electron beams composed of picosecond-long pulses, and of techniques to extract picosecond pulses of X-rays from a synchrotron source, have proved crucial. These pulses can be used for electron or X-ray diffraction studies, or in techniques such as 'X-ray absorption near-edge structure' (XANES) and 'extended X-ray absorption fine structure' (EXAFS), to trace the evolution of atomic positions during a chemical reaction.

Electrons interact strongly with matter, so an electron-diffraction signal is usually easy to detect. However, because of that strong interaction, time-dependent electron diffraction can only be used to study the reactions of isolated molecules, or reactions on an exposed surface. The interaction of X-rays with matter is, in general, much weaker than that of electrons. Time-dependent XANES, EXAFS and X-ray diffraction can be

4. Han, J. S., Szak, S. T. & Boeke, J. D. *Nature* **429**, 268–274 (2004).
5. Han, J. S. & Boeke, J. D. *Nature* **429**, 314–318 (2004).
6. Perepelitsa-Belancio, V. & Deininger, P. *Nature Genet.* **35**, 363–366 (2003).
7. Ostertag, E. M. *et al.* *Nature Genet.* **32**, 655–660 (2002).
8. Orgel, L. E. & Crick, F. H. C. *Nature* **284**, 604–607 (1980).

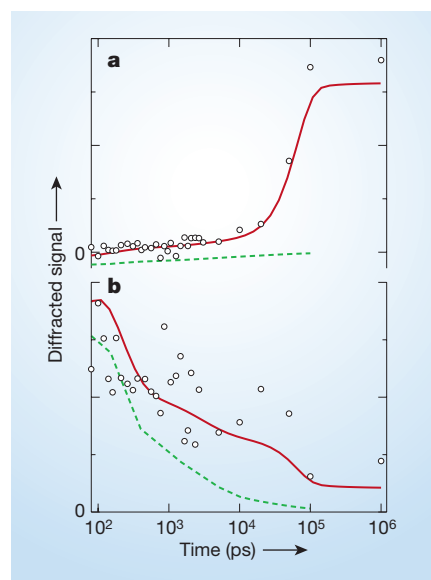


Figure 1 X-ray diffraction of iodine molecules in carbon tetrachloride solution. Plech *et al.* excited the iodine molecules with a laser pulse, then monitored the molecules' dissociation using X-ray pulses applied at varying time intervals after the excitation. The diffracted signal is shown here for two different spatial scales: a, 0.6 nm and b, 0.15 nm. The rise of the data in a is attributed to the solvent taking up the energy released by the iodine molecules as they relax to their ground state, and expanding. The behaviour is quite different in b: on this scale, both the relaxation of the iodine molecules and the recombination of dissociated atoms contribute to the signal. In both cases, the red curve indicates the predicted signal, taking all the relevant effects into account; the green curve represents the contribution from iodine atoms that do not escape their solvent cage.

carried out in condensed phases. Although XANES and EXAFS measurements provide less precise information on atomic structure than does X-ray diffraction, when applied to solutions they avoid many problems associated with background scattering from the solvent. In X-ray diffraction, the occurrence of scattering from the support medium must be taken into account.

Among the first applications of time-resolved X-ray diffraction to a chemical reaction was the study of the structural changes that are associated with the photodissociation of the CO group in myoglobin, traced at the nanosecond timescale. Plech *et al.* have extended the diffraction technique to reach the picosecond timescale and track the photodissociation of iodine molecules dissolved in carbon tetrachloride. This reaction has been examined extensively as a model of the competition between two modes of recombination for the iodine atoms: whether the photodissociated atoms are trapped by the solvent and recombine with each other (called geminate recombination); or whether the iodine atoms escape their cage of solvent and recombine with other iodine atoms (non-geminate recombination).

Plech *et al.* used 150-fs laser pulses to excite the iodine solute, generating a mixture of electronic states that mainly dissociate to ground-state atoms. The excited iodine molecules that are formed have a bond length of 0.27 nm, which stretches to 0.4 nm in less than a picosecond. Next, the molecule either dissociates into two iodine atoms or

loses energy to the solvent, relaxing to an excited bound state on a timescale of 2.7 ns or to the ground state in about 180 ps. By sending intense 100-ps pulses of X-rays into the iodine solution at varying time intervals after the laser excitation pulse, Plech *et al.* mapped changes in the bond lengths between iodine atoms and between neighbouring solvent molecules during the course of the reaction, through the series of diffraction patterns created.

On the question of geminate recombination, Plech *et al.* find that for only 14% of excited iodine molecules do the dissociated atoms escape the solvent cage to recombine with an atom other than their original partner. But a more striking feature of their results is the signature of structural variation in the solvent during the reaction (Fig. 1). When the reaction is probed on a spatial scale of around 0.6 nm, the solvent structure is seen to change slowly for about 20 ns, and then to change rapidly over a period of about 50 ns. But when the reaction is probed on a smaller spatial scale of 0.15 nm, the solvent structure changes at about the same rate for several hundred nanoseconds. The

surprising implication is that although the initial change in solvent structure, arising from the deposition of energy from the recombining iodine atoms, occurs in less than 200 ps, the transformation of that initial change to bulk thermal expansion does not occur for another 10,000 ps.

It is interesting to speculate that the solvent-structure change is a monitor of a coherent excitation that later decays, but this remains to be determined. If, indeed, the motion of an excited molecule and its surrounding shell of solvent molecules is coherent for a reasonable fraction of the reaction time, then our picture of how solute and solvent interact during a reaction will need to be modified. In this case, the solvent is much more than just a source of random energy exchanges with the reactant. Rather, solute and solvent act as a supermolecule, defining the reaction rate and, possibly, also the reaction mechanism.

Stuart A. Rice is in the Department of Chemistry and the James Franck Institute, University of Chicago, 5640 S Ellis Avenue, Chicago, Illinois 60637, USA.
e-mail: sarice@uchicago.edu

Solar System

Captain Cook and the black drop

At intervals, Venus passes directly between the Earth and the Sun, making a transit across the Sun's bright disk. On 8 June, an entire transit of Venus will be visible from most of Asia, Africa and Europe. Australia will experience the transit at around sunset, and across the eastern side of the Americas it will occur around sunrise.

Johannes Kepler, as he plotted the motions of the heavens, was the first to realize that such transits would occur. However, his predicted transit of 1631 was not visible in Europe. The first recorded sighting of a transit was made, in north-west England, by Jeremiah Horrocks, whose own calculations had pinpointed a repeat of the spectacle on 4 December 1639. The intervals between transits follow an unusual pattern: 8 years, then 121.5, then 8, then 105.5 years, and will do so until 2984.

Seventeenth-century astronomers hoped that observations of a transit of Venus would allow them to obtain an accurate measure of the distance from the Earth to the Sun. The concept was eventually formally published in 1716 by Edmund Halley, who predicted from where

the transit of 1761 should be viewed so that measurements of the transit times could be used to calculate the Sun's distance. But his predictions were somewhat awry: Halley mistakenly used a plus sign in his calculation where there should have been a minus.

Mathematical fallibility aside, in fact the transit of Venus has never yielded an accurate measure of the Earth–Sun distance. The calculation is hampered by what is known as the 'black drop problem'. The exact moment at which the dark disk of Venus is fully on the disk of the Sun is hard to ascertain, because the planet's disk seems to stay connected to the dark edge of the Sun: the disk appears as a black drop rather than a perfect circle. Only last year was this effect attributed to the way in which the Sun's brightness lessens at its visible edge, in combination with the natural blurring caused by a telescope.

The difficulty of timing a transit was also noted by Captain James Cook (pictured), on his first round-the-world voyage on the *Endeavour*. In 1769, before reaching the east coast of Australia, Cook was at Tahiti. On Saturday 3 June, he wrote



in his logbook: "This day proved as favourable to our purpose as we could wish. Not a Cloud was to be seen the whole day, and the Air was perfectly Clear, so that we had every advantage we could desire in observing the whole of the Passage of the planet Venus over the Sun's Disk. We very distinctly saw an Atmosphere or Dusky shade round the body of the planet, which very much disturbed the times of the Contact, particularly the two internal ones. Dr. Solander observed as well as Mr. Green and myself, and we differ'd from one another in Observing the times of the Contact much more than could be expected."

By 1882, transit fever had

captured both the public imagination and the newspaper headlines. The *San Francisco Chronicle* reported that "Many of the residents of San Francisco were noticed yesterday with a piece of smoked glass to their eye, looking curiously at the sun". William Harkness, then director of the US Naval Observatory in Washington DC, had warned on the eve of that transit that "there will be no other till the twenty-first century of our era has dawned upon the Earth, and the June flowers are blooming in 2004".

The transit begins shortly after 5:00 Universal Time on 8 June. To view it, the simplest of pinhole cameras will suffice — weather permitting, of course. Alison Wright

Obituary

John Maynard Smith (1920–2004)

John Maynard Smith, one of the greatest thinkers in evolutionary biology, died on 19 April: at the time he was sitting in his chair at home, surrounded by papers over which he had been brooding. He was a generous and extremely gifted man, whose strength was to tackle biological problems with simple mathematical models and to draw insightful conclusions from them; indeed, Maynard Smith showed an aversion to the pompous use of mathematics and the mystification of science that often goes with that. He made crucial contributions to several debates in evolutionary theory: the levels — genes, the individual organism, and so on — at which natural selection operates most effectively; the maintenance of sex as a costly mode of reproduction; the use of game theory for biological analysis of conflict and cooperation; the characteristics of major evolutionary transitions, such as the origin of multicellularity; and the logic of animal signalling.

Maynard Smith was born in 1920, into a British upper-class family with a limited interest in the scientific and cultural matters that were to become so important to him. In due course he was sent to Eton College, an establishment that features near the top of the pecking order of British 'public' (that is, private) schools. He did not have fond memories of his time at Eton. But he did receive a good grounding in mathematics there, which he found very useful for "doing his sums" in his later callings of engineering and theoretical biology. And, as he pointed out in his typically fair way, it was also to Eton's credit that the library included books by his future mentor, J. B. S. Haldane, even though this co-founder of population genetics and vocal member of the Communist Party was held in particular disregard there.

The family preference was for Maynard Smith to become a stockbroker. He firmly rejected this option, but he needed an alternative. Prompted by an earlier visit to Eton by one of the designers of Sydney Harbour Bridge, whom he thought very impressive, he decided to become an engineer, finishing a degree in the subject at Trinity College, Cambridge, in 1941. He was involved in aircraft design during the Second World War, but in 1947 left this occupation partly because his bad eyesight prevented him from pursuing his ambition of flying. He then began his second, breathtaking, career in evolutionary biology, studying at University College



From Eton to Sussex, from aircraft engineer to original thinker in evolutionary biology

London under the guidance of the man who had already inspired him from a distance: the charismatic J. B. S. Haldane, a towering figure in British science at that time, and still a prominent supporter of the Communist Party.

Both men adhered to the Communist ideology for quite some time. Maynard Smith's main motivations for joining the party were its stance on racism and sexism, its fight against social inequality, and its firm anti-Nazi position. He felt that these attributes were synergistic, but in later reflections admitted his blindness to the malign aspects of Communism — to which, for instance, the eyes of George Orwell and Arthur Koestler had been opened earlier because of their more worldly experiences. It was the detrimental influence on Soviet biology of Trofim Lysenko's politically motivated genetics theories that caused Maynard Smith to abandon his overly idealistic view of Communism. He ultimately withdrew from political activism, and left the Communist Party, when Hungary was invaded by Soviet troops in 1956.

Maynard Smith's initial scientific work, developed before Haldane's departure for India in 1957, largely involved experimental work on inbreeding and ageing of the fruitfly *Drosophila*. His most important theoretical insight into ageing was that natural selection would be expected to synchronize the ageing processes of the various subsystems of

an organism. To get a feel for this insight, suppose for a moment that there was no such synchrony and that a 'lead mechanism' of ageing existed instead. Investment in the higher durability of the other subsystems would then be superfluous because they would last longer than needed. Natural selection would counteract such a waste of resources. A similar consideration must have guided Henry Ford to the junk yard of old automobiles, to decide which parts to produce in lower quality in the future, so as to synchronize the expected time at which different parts would wear out.

Maynard Smith was not keen on administrative affairs. But in 1965 he was delighted to accept a proposal, initiated by Peter Medawar, to become the founding dean of the School of Biological Sciences at the University of Sussex, where he stayed happily for the rest of his life. In his hands, Sussex became a respected centre in various branches of biology, and now boasts a John Maynard Smith Building (where the man himself would attend lectures).

Maynard Smith often admitted that he was not a champion chemist, but he was comfortable with molecular biology, where the processing of information plays a central role. His formulation of the concept of 'sequence space' in the late 1960s was visionary. He wanted to know how natural selection could guide proteins through a highly multidimensional space (where proteins differing in only one amino acid would be nearest neighbours), to assume various distributions of functionality in that space ('continents', 'islands', and so on). He was about 20 years ahead of his time with these early considerations, which are now widely applied in molecular evolution both at the protein and the RNA level.

In 1971, he formulated a basic problem in evolutionary biology known as the 'twofold cost of sex'. It goes as follows: other things being equal, a population of parthenogenic females — females that reproduce without fertilization — should grow at a rate twice that of sexually reproducing females, because half of the offspring of the latter are males. With admirable frankness, he later called attention to the fact that the problem was seen in much the same way by the German biologist August Weismann — whom, after distancing himself from his previous prejudices against the man, he considered to have been second only to Charles Darwin as a nineteenth-century thinker about evolution. The characteristically lucid posing of the problem has stimulated much research on the evolutionary maintenance of sex — research in which both theoreticians and empiricists have

asked what advantages could compensate for the twofold cost.

It is fair to wonder about the common thread behind Maynard Smith's various lines of investigation. From discussions with him, we think that it was the application of rigorous evolutionary reasoning to phenomena that at first sight seem to contradict darwinian theory. For example, if organisms should maximize their 'fitness', then why do they age? Dead organisms do not reproduce; hence their fitness does not increase any more. The cost of sex obviously poses a similar problem, and in the same vein it is a puzzle why so many antagonistic interactions between animals of the same species seem to be ritualized encounters rather than serious fights. Competing animals should try to attack and kill one another, shouldn't they? There are two proposed answers to this question. The first is that it is good for the species to avoid harmful encounters by finding ritualized ways of resolving conflict. Maynard Smith, however, argued convincingly that a second type of answer is the correct one. The idea, which he worked out together with George Price, is that ritualized contests can be advantageous to the individual by allowing an assessment of the probable outcome of fighting. This way, the weaker individual can escape or surrender before the opponent escalates the contest; the stronger individual avoids possible injury. Overly aggressive animals foregoing a ritual display would too often be harmed by opponents who fought back.

In their analysis of animal behaviour during contests, Maynard Smith and Price laid the foundation for the new field of evolutionary game theory. Traditionally, the theory of games dealt with rational human decision-making in situations where there is a conflict of interest among interacting individuals. The new branch of game theory was one in which natural selection acts instead as the decision maker. The framework of evolutionary game theory has turned out to be extremely useful, and has been generalized to plants and even to microbes and replicating molecules. At the heart of problems in evolutionary game theory lies a complex optimization problem, where an individual's best 'strategy' depends on what strategies other members of the population are using. A trivial example is driving on the left side of the road. In the population where Maynard Smith tended to use his car, it is advisable to drive on the left side — not because this is good in itself but because, as all other the drivers are doing it, it is the best strategy for avoiding an accident.

Evolutionary game theory considers



genes and individuals as units among which there may be conflict. But it was also clear to Maynard Smith that there can be a problem with overemphasizing these units and their competition. In evolution, higher-level units (such as a plant or animal cell) repeatedly formed from lower-level ones (such as bacteria). If competitive evolution at the lower level had not been suppressed somehow, the higher-level units would have been disrupted. How is this possible? A thorough investigation of major transitions in evolution revealed some common themes: multi-level selection, origin of novel inheritance systems, combination of function and division of labour, and so on. We have yet to reach a full understanding of any of the major transitions, but many scientists have been stimulated to enter this field.

Asking a new question is sometimes as useful as answering an old one. Maynard Smith observed succinctly that twentieth-century biology is more about the role of information in biology than about anything else. The terminology of molecular biology (transcription, translation, proofreading and so on), the concept of positional information in embryology, the nervous system as information processor and the questions on animal signalling (about which he published his last book just a few months ago) all confirm the validity of his point.

Although it was not always apparent, Maynard Smith had been a naturalist since his childhood. Following the tradition of the British gentleman as a keen observer of the countryside, he was convinced that science must be deeply grounded in empiricism. He would always start with a concrete problem and boil it down to the key issues. Of course, his engineering

past occasionally surfaced. For example, his models of a dual inheritance system (genetics and epigenetics) have a strong engineering flavour. They were published in the *Journal of Theoretical Biology* — this was a journal close to his heart, and he served on its editorial board almost until his death.

John Maynard Smith — widely known as JMS — was an exceptionally lucid thinker and a delightful man. He would talk about science to anyone and, in the same open-minded way, treat young and established people with equal respect. This was apparent at an international meeting on evolution and systematics at Sussex in 1986 (shortly after his cancer operation), when he listened during a long break to a completely unknown foreign youngster explaining a novel model of early evolution, having chased everybody else away, including celebrities of the time. And he was willing to give up preconceptions: his reappraisal of Weismann is a case in point.

Some of us will find it difficult to live without him. He said that he always wanted to follow in the footsteps of Haldane, and in many respects he was successful in that aim. Emulating John Maynard Smith, mentor, colleague and man, will be an immense challenge. **Eörs Szathmáry and Peter Hammerstein**
Eörs Szathmáry is at the Collegium Budapest, Szentháromság utca 2, Budapest, and Eötvös University, Hungary. Peter Hammerstein is at the Institute for Theoretical Biology, Humboldt University Berlin, Invalidenstrasse 43, 10115 Berlin, Germany.
e-mails: szathmary@colbud.hu
p.hammerstein@biologie.hu-berlin.de

See also: www.nature.com/nature/focus/maynardsmith

Aquatic ecology

Growth prospects for algae

Ecol. Lett. doi:10.1111/j.1461-0248.2004.00597.x (2004)

The amount of CO₂ in the atmosphere could double over the coming century. How will that affect photosynthesis by the world's phytoplankton?

Peter Schippers *et al.* modelled the uptake of carbon by five marine and two freshwater species of alga. They calculated the organisms' rates of carbon uptake at today's atmospheric CO₂ concentration and at an increased concentration, taking into account the effect this uptake would have on the carbon chemistry of the surrounding water and the subsequent availability of carbon in the form of dissolved CO₂ and bicarbonate ions. They found that when algae take up large amounts of carbon, the pH of the water increases sharply, the water CO₂ concentration falls, and the amount of CO₂ entering the water becomes proportional to the concentration of CO₂ in the atmosphere.

From their results, and assuming nutrients are not limiting, Schippers *et al.* conclude that carbon uptake by phytoplankton could double with a doubling of atmospheric CO₂ — and that the incidence of harmful algal blooms would increase.

Lucy Odling-Smee

Astronomy

Illuminating the dark age

Preprint at <http://arXiv.org/abs/astro-ph/0312134> (2004)

After the Big Bang, but before the first stars ignited, the Universe was shrouded in darkness. Quasars — among the brightest and most distant objects known — illuminated the Universe when it was less than a billion years old. Their light bears the imprint of absorption by neutral hydrogen gas in the early Universe.

Abraham Loeb and Matias Zaldarriaga propose that there may be a similar imprint of the even earlier Universe, borne by the cosmic background radiation, that captures the depths of the cosmic dark age. When the Universe was between 20 million and 100 million years old, neutral hydrogen gas would have resonantly absorbed cosmic background radiation, the residual echo of the Big Bang. If this absorption pattern could be mapped across the sky, it would reveal how the neutral hydrogen was distributed in the dark-age Universe.

Today, say Loeb and Zaldarriaga, that information is to be found at radio wavelengths, from 6 to 21 metres, which are difficult to observe using current technology. The next generation of telescopes might, however, catch these radio waves, and illuminate the dark age.

Mark Peplow



Plant genetics

Grape expectations

Science **304**, 982 (2004)

Black and red grapes owe their colour to red pigments — anthocyanins — in their skins. According to one account, the white varieties arose as the plants producing red-skinned fruit picked up pigment-altering mutations. From their investigations, however, Shozo Kobayashi *et al.* hypothesize that black grapes gave rise to white grapes, which in turn yielded red varieties.

The team analysed the genetic make-up of different coloured varieties and found that white grapes contain a string of intruder DNA, called a retrotransposon, inside a gene that normally promotes pigment production. As a result, the gene's expression is blocked and the grape becomes white rather than black. Red grapes developed from white varieties as they lost the retrotransposon, picking up a second mutation that switched pigment production back on.

Kobayashi and colleagues argue that mutations in a single gene determine skin colour in grapes, and suggest that the same mutant gene has spread among most, if not all, white varieties in the world.

Helen R. Pilcher

Physics

A quantum transistor

Phys. Rev. Lett. **92**, 176801 (2004)

Transistors are an integral part of logic gates, the basic electronic components of computer circuits. A quantum transistor could potentially process quantum bits in which data may exist in both 'on' and 'off' states at the same time, thus multiplying computing power.

J. C. Chen *et al.* have made such a quantum transistor from two tiny discs of gallium arsenide. The discs, or 'quantum dots', have a diameter of about 200 nanometres, and are kept about 100 nanometres apart on an insulating wafer. Each holds a pool of 40 to 60 electrons, fed by an external circuit that delivers electrons one by one into the dots.

In experiments performed at 30 mK, Chen *et al.* added a single excess electron to each dot, the (unpaired) spin of that electron determining the spin of the dot itself — either 'up' or 'down'. Both dots were set up with matching spins. Then, by varying the gate voltage across the double-dot system, the authors tuned the interaction between the dots, such that their spins became entangled. This mixed configuration is both 'spin-up' and 'spin-down' — and so could be the basis of an electronic switch that is both 'on' and 'off'.

Mark Peplow

Stem cells

Damage alert

J. Clin. Invest. **113**, 1364–1374 (2004)

Neural stem cells in mammalian brains are known to migrate to injured brain regions, perhaps to carry out repair. In a step towards potentially using these cells therapeutically, Lixin Sun *et al.* have identified a molecule that lures them to their destination.

The authors searched for genes that are expressed at higher levels in injured mouse brains than in healthy ones. From nearly 200 candidates, they homed in on one gene, which encodes stem cell factor (SCF) — a secreted protein that has already been linked to the division and migration of other stem-cell types.

Sun *et al.* confirmed that SCF is produced by damaged neurons and that the receptor for SCF, c-kit, is expressed on neural stem cells. They also showed that labelled stem cells move towards SCF after it is injected into a mouse's brain. The authors speculate that SCF might one day be used to boost migration of stem cells to sites of disease or damage in the brain or spinal cord. There, the cells might deliver therapeutic molecules, or be converted into new tissue by other chemical signals.

Helen Pearson

Red deer stocks in the Highlands of Scotland

A drastic culling of deer may not be the best strategy to arrest erosion of heather cover.

Grazing by hill sheep and red deer prevents the regeneration of woodland in many parts of the Scottish Highlands and has also led to extensive loss of heather cover^{1–3}. Conservation bodies claim that there has been a rapid rise in Highland deer numbers caused by inadequate management and that these need to be drastically reduced⁴. Here we show that the recent increase in red deer stocks has probably been overestimated and suggest that the gradual rise in numbers since 1970 may be a consequence of a reduction in sheep stocks and of changes in winter weather, rather than of a reduction in culling rate. Although there would be environmental benefits in reducing deer numbers, there is an equal need to reduce the numbers of hill sheep in many parts of the Highlands.

In the British Isles, few environmental issues have generated such a long-running debate as the management of the Scottish Highlands, which have been grazed by a combination of red deer and sheep for much of the past two hundred years^{1,5,6}. During this period, heather cover has diminished and the prevention of tree regeneration by browsing has led to the gradual erosion of many of the remaining native woodlands^{2,3}.

A recent report commissioned by the WWF (formerly the World Wildlife Fund) and the Royal Society for the Protection of Birds (RSPB)⁴ claims that red deer (*Cervus elaphus*) in the Scottish Highlands increased from 300,000 in 1989 to 450,000 by 2002 (Fig. 1a). The report predicts dire environmental consequences unless deer numbers are drastically reduced. However, we believe that the concerns of conservation bodies may be exaggerated.

First, it is unlikely that Scottish red deer numbers are increasing as rapidly as the report claims⁴. A large part of the apparent increase is due to the arbitrary tripling of the estimated number of deer living in woodland, where they cannot be counted reliably (from 30,000 in 1990 to 100,000 in 2002; Fig. 1a). When figures are restricted to deer living on the open hill (where numbers can be estimated more reliably) and are corrected for the area and year in which counts were made, our analysis indicates that the increase

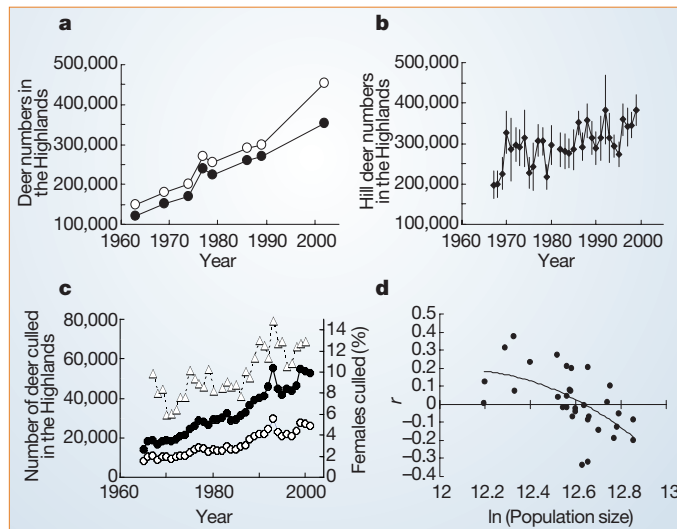


Figure 1 Estimates of changes in red deer populations in the Scottish Highlands since 1960. **a**, WWF/RSPB estimates of total red deer in Scotland⁴, including (hollow circles) or excluding (filled circles) an arbitrary increase in the estimated number of woodland deer by 70,000 in 2002 (from ref. 4); **b**, total hill deer numbers estimated from a multiple regression model⁶ and corrected for year of count, showing 95% confidence limits; **c**, estimates of annual numbers of adult females (hollow circles) and all deer (filled circles) culled and of the annual culling rate (per cent) of adult females (triangles); **d**, estimates of annual population-growth rate (r) of hill deer, plotted against population size in the following year, using data shown in **b**. The plot indicates that population growth is density-dependent and that $r = 0$ at a population size of 310,000–315,000, given the existing culling regime.

in total deer numbers and average density since 1970 has probably been gradual (Fig. 1b); also, it is unclear whether the rate of growth has accelerated in recent years. Moreover, the confidence interval for any realistic estimate of total hill deer numbers is very large: the report⁴ provides a single estimate for 2002 without error range (354,000) that does not fall far outside the upper confidence limit of the total estimated in the 1980s (337,000; ref. 6).

Second, it is unlikely that recent increases in deer numbers have been caused by a relaxation of management, because (despite some fluctuations between years) annual culls have been substantially higher since 1985 than they were before this (Fig. 1c). Instead, the gradual increase in hill deer numbers since 1970 may reflect continuing reductions in the number of hill sheep and trends in weather conditions that have led to less severe winters and warmer summers, which have previously been shown to be associated with changes in hill deer numbers^{6,7}. In fact, deer populations in the Highlands may have been resource-limited for some time: between 1960 and 1990, annual culls of females in many parts of the Highlands were below the 15–20% level that is necessary to prevent an increase in population size^{8,9} and there has been no lasting reduction in total deer numbers since then (Fig. 1b). Moreover,

annual changes in population-growth rate (r) values since 1960 are probably density-dependent, with $r = 0$ at a total population size of 310,000–315,000 deer under the current culling regime (Fig. 1d).

Third, although even moderate deer densities can prevent woodland regeneration and damage forestry plantations^{2,3,9,10}, there is limited evidence that deer populations cause extensive changes in the distribution of other upland ecosystems⁸, unless numbers are sustained at a high level by the provision of supplementary food in winter. In many areas grazed by deer alone, community boundaries appear to be stable^{9,11}. Extensive heather loss (one of the principal concerns of the WWF and RSPB) might be more frequently caused by sheep (which are often maintained at summer densities that are several times higher than deer) or by a combination of sheep and deer^{6,9}, than by deer alone.

So what should be done to protect the environment of the Scottish Highlands? The RSPB have shown that it is possible to encourage woodland regeneration and arrest heather loss by reducing grazing pressure to very low levels¹⁰, but reducing sheep and deer numbers to a comparable extent across the Highlands would affect many human communities. A more feasible solution could be to reduce the combined stocking rate of sheep and deer to a level at which heather loss is arrested and to protect sensitive areas of woodland with fences (despite their disadvantages) to allow regeneration to continue. Even this aim will be difficult to implement as sustained control of extensive populations is seldom easy or cheap^{12,13}.

T. H. Clutton-Brock*, T. Coulson†,

J. M. Milner‡

*Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK
e-mail: thcb@cam.ac.uk

†Department of Biological Sciences, Imperial College at Silwood Park, Berkshire SL5 7PY, UK

‡Hedmark University College, Evenstad, 2480 Koppang, Norway

1. Mitchell, B., Staines, B. W. & Welch, D. *Ecology of Red Deer: A Research Review Relevant to their Management in Scotland* (Institute of Terrestrial Ecology, Cambridge, 1977).

2. Beaumont, D., Dugan, D., Evans, G. & Taylor, S. in *Our Pinewood Heritage* (ed. Aldhous, J. R.) 73–95 (Forestry Commission, Edinburgh, 1994).

3. Milne, J. A., Birch, C. P. D., Hester, A. J., Armstrong, H. &

- Robinson, A. *The Impact of Vertebrate Herbivores on the Natural Vegetation of the Scottish Upland: A Review* (Scottish Natural Heritage, Edinburgh, 1998).
4. Hunt, J. F. *Impacts of Wild Deer in Scotland: How Fares the Public Interest?* (Report for WWF Scotland and RSPB Scotland, 2003).
5. Hart-Davis, D. *Monarch of the Glen* (Cape, London, 1978).
6. Clutton-Brock, T. H. & Albon, S. D. *Red Deer in the Highlands* (Blackwell Scientific, Oxford, 1989).
7. Clutton-Brock, T. H. & Albon, S. D. *Nature* **358**, 11–12 (1992).
8. Clutton-Brock, T. H., Crawley, M. J. & Milner, J. M. in *A Highland Deer Herd and its Habitat* 237–255 (Red Lion, London, 2002).
9. Milner, J. M., Alexander, J. & Griffin, C. in *A Highland Deer Herd and its Habitat* 109–146 (Red Lion House, London, 2002).
10. Hester, A. J., Iason, G. R., Stolte, A. & Sim, D. *Vegetation Survey: Abernethy Forest RSPB Reserve* (RSPB Reserve, Edinburgh, 1998).
11. Darling, F. F. A *Herd of Red Deer* (Oxford Univ. Press, 1937).
12. Caughley, G. in *Rangelands: A Resource under Siege* (eds Joss, P. J., Lynch, P. W. & Williams, O. B.) 135–140 (Australia Academy of Science, 1986).
13. Caughley, G. in *Problems in Management of Locally Abundant Wild Animals* (eds Jewell, P. A., Holt, S. & Hart, D.) 7–19 (Academic Press, New York, 1981).

Competing financial interests: declared (see online version of this communication for details).

Vision

Steady-state misbinding of colour and motion

When you see a red ball rolling across the floor, the ball's redness, roundness and motion appear to be unified and inseparably bound together as features of the ball. But neurophysiological evidence indicates that visual features such as colour, shape and motion are processed in separate regions of the brain¹. Here we describe an illusion that exploits this separation, causing colour and motion to be recombined incorrectly while a stable stimulus is being viewed continuously.

The illusion is seen in stimuli containing two sheets of random dots, where one sheet is moving up and one is moving down. The sheets contain dots of two colours such that the central and peripheral portions of the stimuli combine colour and motion in opposite fashions (Fig. 1a). On the upward-moving sheet, dots in the centre are red and dots in the periphery are green. On the downward-moving sheet, dots in the centre are green and dots in the periphery are red. (For movie, see supplementary information.)

Observers gazing at the centre of the display perceive peripheral dots erroneously: they 'bind' colour and motion in the wrong combination. The entire display therefore appears to be covered by a sheet of red dots that are moving upwards and a sheet of green dots that are moving downwards.

To quantify this effect, observers were asked to report whether the majority of peripheral red dots were moving up or down over a series of trials in which we varied the percentage of peripheral dots moving in either direction. In control trials, where the central portion of the stimulus was omitted, responses of observers ($n=5$; 3 naive) roughly followed the physical stimulus (Fig. 1b, black line). But when the central

portion was present, responses shifted markedly, and in the same direction as the central dots (Fig. 1b, red and green lines; MANOVA, $F(10,16) = 35.2$, $P < 0.001$).

To ensure that observers' responses were reflecting their perception and not a simple bias to respond to the central dots, we generated stimuli in which one peripheral region matched the centre and the other contained the opposite motion. Observers ($n=4$; 3 naive; all from the first experiment) perceived both sides to move with the centre, performing poorly in reporting which side moved the other way (mean $\pm$ s.e.m.: $47 \pm 8\%$ correct; sensitivity measure $d' = -0.12 \pm 0.38$). On the other hand, they performed well in reporting which side was which when the stimulus centre was blank ($86 \pm 2.5\%$ correct; $d' = 1.5 \pm 0.2$, paired t -test $P < 0.5$).

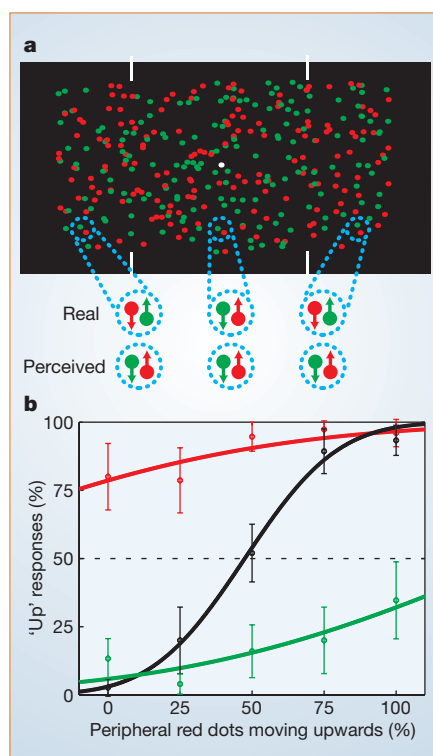


Figure 1 Steady-state misbinding stimulus and observers' reports. **a**, Sample transparent motion stimulus. Two sets of random dots move rigidly in opposite directions. 'Real' circles show details of the physical stimulus. In the upward-moving surface, dots near fixation are red, and dots beyond 6° of horizontal eccentricity are green; in the downward-moving surface, the reverse coloration applies. White bars marked the transition line. 'Perceived' circles show details of the illusory percept. Observers incorrectly pair colour and motion in the periphery. Across the entire stimulus, upward-moving dots appear red and downward-moving dots appear green, forming two homogeneous surfaces of dots moving in opposite directions. **b**, Proportion of trials in which observers judge that most red dots beyond the white bars are moving upwards, plotted as a function of the actual percentage of red dots moving upwards; time to view stimulus was unlimited. In control trials (black line), the central region is blank, and responses follow the stimulus. When the central region contains upward-moving red dots and downward-moving green dots (red line), responses shift predominantly to 'up'. When the central region contains the opposite motion (green line), responses shift predominantly to 'down'.

We propose that the illusion described here is the result of an ambiguity-resolving mechanism. Because vision is clearest at the point of gaze, it would normally be advantageous to use information from the centre to resolve peripheral ambiguity if the physical source is likely to be a common uniform surface. Uniform surfaces are strongly implied by the contiguous and precisely equal values of colour and motion present across this stimulus. Peripheral features are bound in such a way as to accord with the usually more reliable centre, although in this case the binding is mistaken.

Previous investigations of basic visual feature binding required the use of stimuli with brief presentation times or rapidly changing features to induce errors or inefficiencies^{2–4}. The illusion presented here occurs despite fully attentive and continuous viewing, avoiding the common confounds of memory, expectation and task strategy⁵. It solidifies the evidence for the existence of a binding problem and provides a substrate for neurophysiological investigation.

Daw-An Wu*, **Ryota Kanai†**, **Shinsuke Shimojo*‡§**

*Division of Biology and ‡Computation and Neural Systems, California Institute of Technology, 139-74, Pasadena, California 91125, USA
e-mail: daw-an@caltech.edu

†Psychonomics Division, Helmholtz Research Institute, Universiteit Utrecht, 3584 CS Utrecht, The Netherlands

§Human and Information Science Laboratory, NTT Communication Science Laboratories, Atsugi, Kanagawa 243-0198, Japan

- Livingstone, M. & Hubel, D. *Science* **240**, 740–749 (1988).
- Treisman, A. M. & Gelade, G. *Cogn. Psychol.* **12**, 97–136 (1980).
- Moutoussis, K. & Zeki, S. *Proc. R. Soc. Lond. B* **264**, 393–399 (1997).
- Nishida, S. & Johnston, A. *Curr. Biol.* **12**, 359–368 (2002).
- Wolfe, J. M. & Cave, K. R. *Neuron* **24**, 11–17 (1999).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

Corrigendum

Insecticide resistance in mosquito vectors

M. Weill, G. Lutfalla, K. Mogensen, F. Chandre, A. Berthomieu, C. Berticat, N. Pasteur, A. Philips, P. Fort, M. Raymond
Nature **423**, 136–137 (2003).

It has been drawn to *Nature's* attention that M. W., N. P., P. F. and M. R. are named as inventors on two patent applications relevant to this work (FR20020007622 20020620, filed in June 2002; and FR200200213799 20021105, filed in November 2002), which should therefore have been declared as competing financial interests.

brief communications arising online
♦ www.nature.com/bca

Evolutionary biology: Lamprey *Hox* genes and the evolution of jaws

Y. Takio, M. Pasqualetti, S. Kuraku, S. Hirano, F. M. Rijli, S. Kuratani (doi:10.1038/nature02616)

Evolutionary biology

Lamprey *Hox* genes and the evolution of jaws

Arising from: M. J. Cohn *Nature* **416**, 386–387 (2002).

In vertebrates with jaws (gnathostomes), the jaws are formed from the first pharyngeal arch (PA1), which does not express homeobox (*Hox*) genes. Cohn¹ describes expression of the *HoxL6* gene in the PA1 of the lamprey *Lampetra fluviatilis*, a jawless (agnathan) vertebrate, and postulates that a retreat of *Hox* expression from PA1 might have favoured the evolution of jaws in the gnathostome lineage after the split from agnathans¹. Here we examine the distribution of *Hox* genes in another lamprey species, *Lethenteron japonicum*, and find that none are expressed in the PA1. We conclude that Cohn's finding is not a general feature within the lamprey group and is therefore unlikely to be related to jawlessness.

In gnathostome embryos, *Hox* genes display rostral boundaries of expression that are collinear with their relative chromosomal positions^{2,3}. Nested collinear *Hox* expression patterns display sharp anterior boundaries that map to distinct hindbrain segments (rhombomeres) and pharyngeal arches that give rise to the neural-crest-derived skeletal structures in the head and neck region⁴. Notable exceptions are the first rhombomere (r1) and the jaw-forming first pharyngeal arch (PA1), which do not express *Hox* genes. Analysis of *Hox* expression patterns in the jawless lamprey may therefore provide insight into how the gnathostome *Hox* code has been established and its relevance to the evolution of craniofacial structures in vertebrates.

We isolated 11 *Hox* complementary DNAs from embryos of the Japanese lamprey *L. japonicum* (*Lj*). Alignment of the deduced amino-acid sequences of isolated cDNAs, as well as phylogenetic comparisons, revealed that they belong to paralogue groups (PG) 2 to 9, designated *LjHox2*, -3d, -4x, -4w, -5i, -5w, -6w, -6/7m, -8p, -Q8 and -9r, respectively (GenBank accession numbers: AY497314, AB125269–AB125278; Fig. 1a). On the basis of comparison of the nucleotide sequences, *LjHox4x*, *LjHox4w*, *LjHox5w*, *LjHox6w* and *LjHoxQ8* are the

orthologues of *Hox4x*, *Hox4w*, *Hox5w*, *Hox6w* and *HoxQ8*, respectively, in *Petromyzon marinus*^{5–7}, a lamprey species belonging to a different genus. *P. marinus* has three or four *Hox* clusters^{5–7}. Our results from *L. japonicum* are in keeping with this, as we never isolated

more than three *Hox* genes from any single cognate group.

All the *LjHox* genes are strongly expressed at the pharyngula stage. In the neural tube, *LjHox* PG2–8 genes display offset rostral expression borders, moving from anterior to posterior (Fig. 2a–i). The rostral boundary of *LjHox2* expression is in the rhombomere r2 (results not shown), which is similar to that occurring in gnathostomes, where *Hoxa2* is the only *Hox* gene expressed in r2 (refs 4, 8).

The anterior border of *LjHox3d* expression occurs in the middle of r4 (ref. 9). More posteriorly, *LjHox4w*, *LjHox6w* and *LjHoxQ8*, which may be linked in the same cluster^{6,7}, display clear collinear expression patterns (Fig. 2d, f, i). This suggests that collinear *Hox* expression in the embryonic central nervous system represents an ancestral character for vertebrates. It is notable that partial *Hox* collinearity is already present in the neural tube of protochordates¹⁰.

By contrast, in the pharyngeal arches only *LjHox2* and *LjHox3d* display conserved collinear expression patterns in the neural-crest-derived ectomesenchyme (Fig. 2j–m). As in gnathostomes, the rostral expression boundaries of *LjHox2* and *LjHox3d* are present in PA2 and PA3, respectively. *LjHox4x* is instead expressed at low levels, with no clear anterior boundary (Fig. 2c,n), whereas *LjHox4w* and *Hox* PG5–8 are not expressed in the pharyngeal arch ectomesenchyme (Fig. 2d–i), despite their collinear patterns in the neural tube. In the pharyngeal endoderm, *Hox* PG2 and PG3 genes are not expressed, whereas *LjHox4w*, -5i, -6w and -Q8 are all co-expressed in the most caudal pouch (Fig. 2d–f, i–k, o).

None of the *LjHox* genes was expressed in the PA1 at any of the stages that we analysed (Fig. 2, and data not shown), in contrast to the findings of Cohn in *Lampetra fluviatilis*¹. In gnathostomes, the ectopic expression of *Hoxa2* in the *Hox*-free PA1 suppresses jaw development, whereas *Hoxa2* inactivation results in the transformation of PA2 into jaw-related structures^{11–13}. The expression pattern of *LjHox2* is strictly conserved, which is consistent with lamprey PA1 and PA2 being patterned by molecular mechanisms that are similar to those in gnathostomes¹⁴. In agreement with this, *Fgf8*, which is expressed

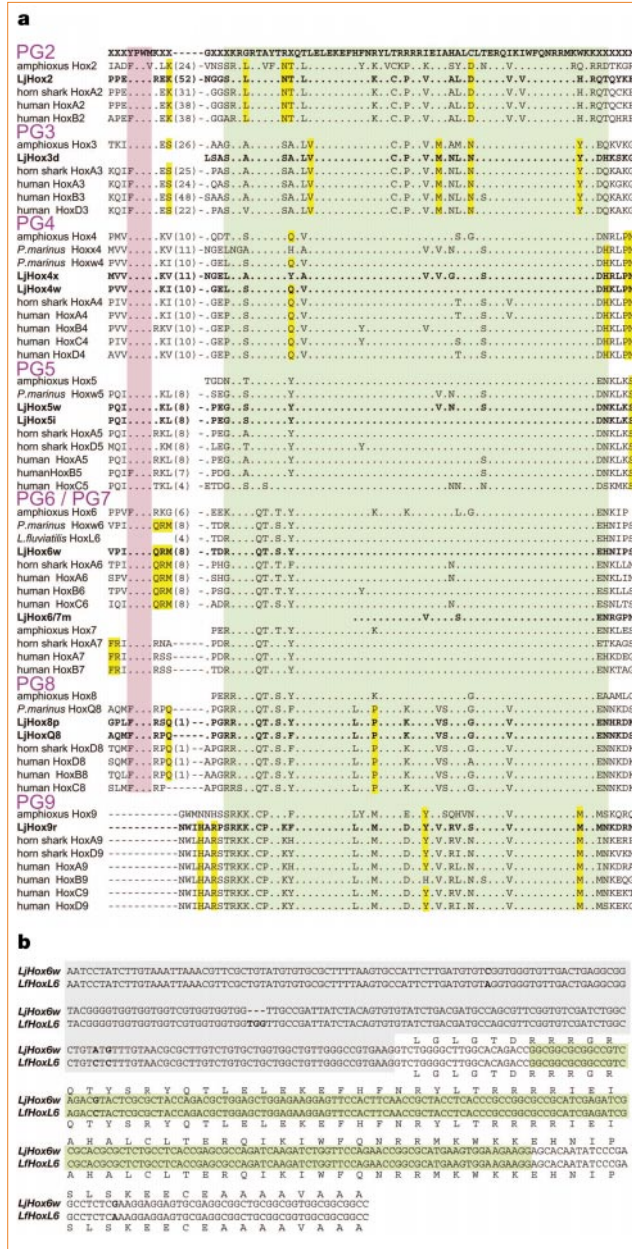


Figure 1 Identification of *Lethenteron japonicum* (*Lj*) *Hox* genes. **a**, Alignment of deduced amino-acid sequences of *LjHox* genes with those of other chordates and the assignment of paralogue groups (PGs) are shown. Homeodomains are indicated in green and the YPWM motif, which is unique to *Hox* PG1–8 genes, in pink. *LjHox* sequences are shown in bold. Amino acids characteristic of each paralogue group are in yellow. **b**, Alignment of *LjHox6w* and *Lampetra fluviatilis* (*Lf*) *HoxL6* nucleotide sequences with their deduced amino-acid sequences. Intronic regions are shaded; homeodomains are in green. Substituted nucleotide sites between the two sequences are in bold. Note that the intronic regions have accumulated a few changes and that the only two substitutions in the exonic region are synonymous.

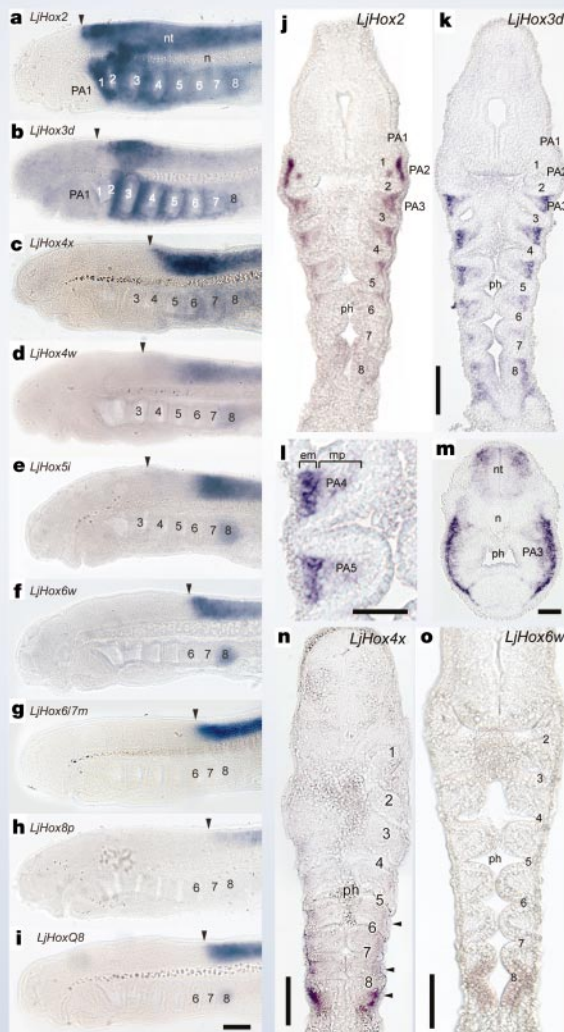


Figure 2 Expression patterns of *Hox* genes in *Lethenteron japonicum* (*Lj*) as revealed by *in situ* hybridization and whole-mount immunostaining. **a–i**, Expression of *LjHox* genes in stage-26 larvae. Arrowheads indicate rostral expression boundaries. **j, k**, Expression of *LjHox2* (**j**) and *LjHox3d* (**k**) in horizontal sections of stage-25 embryos. **l**, High-magnification micrograph of pharyngeal arches (PAs). Positively staining cells correspond to the neural-crest-derived ectomesenchyme (em). **m**, *LjHox3d*-positive neural crest cells populate PA3. **n, o**, Expression of *LjHox4x* (**n**) in the posterior pharyngeal ectomesenchyme (arrowheads) and *LjHox6w* (**o**) in the endodermal pharyngeal pouch 8. Key: mp, muscle plate; n, notochord; nt, neural tube; ph, pharynx; 1–8, pharyngeal pouches. Scale bars: **a–i**, 0.1 mm; **j, k**, 0.1 mm; **l, m**, 0.05 mm; **n, o**, 0.1 mm. Full methodological details are available from the authors.

at the mid–hindbrain boundary in gnathostomes and prevents *Hox* expression in PA1 (ref. 15), is also expressed at the lamprey mid–hindbrain boundary¹⁴. Moreover, the nucleotide sequences of *LjHox6w* and *HoxL6* (ref. 1) display a high degree of homology, indicating that these genes are orthologous (Fig. 1b).

The discrepancy between Cohn's analysis and ours remains unclear. One possibility is that it might be due to genus- or species-specific regulatory differences in lamprey *Hox6* expression patterns. In any event, the *Hox6* expression in PA1 observed in *L. fluviatilis* does not appear to be a general feature within the lamprey group and is therefore not functionally relevant to jawlessness. Rather, the lack of *Hox* expression in the lamprey PA1 may reflect the fact that in both lampreys and gnathostomes the rostral-most pharyngeal arch forms highly specialized structures that are morphologically distinct from those of more posterior arches.

Yoko Takio*†, Massimo Pasqualetti‡§, Shigehiro Kuraku*, Shigeki Hirano¶, Filippo M. Rijli‡, Shigeru Kuratani*

*Laboratory for Evolutionary Morphology, Center for Developmental Biology, RIKEN Kobe, 2-2-3 Minatogima-minami, Kobe 650-0047, Japan
e-mail: saizo@cdb.riken.go.jp

†Graduate School of Science and Technology, Kobe University, Kobe 657-8501, Japan

‡Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, CU de Strasbourg, BP 10142, 67404 Illkirch Cedex, France

§Laboratori di Biologia Cellulare e dello Sviluppo, Università di Pisa, 56010 Pisa, Italy

¶Department of Medical Technology, School of Health Sciences, Faculty of Medicine, Niigata University, Niigata 951-8518, Japan

doi:10.1038/nature02616

1. Cohn, M. J. *Nature* **416**, 386–387 (2002).
2. Duboule, D. & Dollé, P. *EMBO J.* **8**, 1497–1505 (1989).
3. Graham, A., Papalopulu, N. & Krumlauf, R. *Cell* **57**, 367–378 (1989).
4. Hunt, P. et al. *Nature* **353**, 861–864 (1991).
5. Amores, A. et al. *Science* **282**, 1711–1714 (1998).
6. Force, A., Amores, A. & Postlethwait, J. H. *J. Exp. Zool.* **294**, 30–46 (2002).
7. Irvine, S. Q. et al. *J. Exp. Zool.* **294**, 47–62 (2002).
8. Prince, V. & Lumsden, A. *Development* **120**, 911–923 (1994).
9. Murakami, Y. et al. *Development* **131**, 983–995 (2004).
10. Wada, H., Garcia-Fernandez, J. & Holland, P. W. H. *Dev. Biol.* **213**, 131–141 (1999).
11. Rijli, F. M. et al. *Cell* **75**, 1333–1349 (1993).
12. Pasqualetti, M., Ori, M., Nardi, I. & Rijli, F. M. *Development* **127**, 5367–5378 (2000).
13. Couly, G., Grapin-Botton, A., Coltey, P., Ruhin, B. & Le Douarin, N. M. *Development* **125**, 3445–3459 (1998).
14. Shigetani, Y. et al. *Science* **296**, 1316–1319 (2002).
15. Trainor, P. A., Ariza-McNaughton, L. & Krumlauf, R. *Science* **295**, 1288–1291 (2002).

Regional climate shifts caused by gradual global cooling in the Pliocene epoch

Ana Christina Ravelo¹, Dyke H. Andreasen^{2*}, Mitchell Lyle³, Annette Olivarez Lyle³ & Michael W. Wara¹

¹Ocean Sciences Department, and ²Earth Sciences Department, University of California, Santa Cruz, California 95064, USA

³Center for Geophysical Investigation of the Shallow Subsurface, Boise State University, 1910 University Drive, Boise, Idaho 83725, USA

* Present address: Rutgers University, Institute of Marine and Coastal Sciences, 71 Dudley Road New Brunswick, New Jersey 08904, USA

The Earth's climate has undergone a global transition over the past four million years, from warm conditions with global surface temperatures about 3 °C warmer than today, smaller ice sheets and higher sea levels to the current cooler conditions. Tectonic changes and their influence on ocean heat transport have been suggested as forcing factors for that transition, including the onset of significant Northern Hemisphere glaciation ~2.75 million years ago, but the ultimate causes for the climatic changes are still under debate. Here we compare climate records from high latitudes, subtropical regions and the tropics, indicating that the onset of large glacial/interglacial cycles did not coincide with a specific climate reorganization event at lower latitudes. The regional differences in the timing of cooling imply that global cooling was a gradual process, rather than the response to a single threshold or episodic event as previously suggested. We also find that high-latitude climate sensitivity to variations in solar heating increased gradually, culminating after cool tropical and subtropical upwelling conditions were established two million years ago. Our results suggest that mean low-latitude climate conditions can significantly influence global climate feedbacks.

Global climate change of the past 4 million years (Myr) includes the end of the early Pliocene warm period (5–3 Myr ago) and significant intensification of Northern Hemisphere glaciation (NHG) ~2.75 Myr ago (Fig. 1a). The amplitude of 10⁴–10⁶-year climate oscillations increased as climate cooled. The past 4 Myr, unlike the more recent past, can be studied to assess climate theories that: involve climate components with relatively long timescales of response (for example, deep ocean, cryosphere), predict different behaviour in warm versus cold conditions, and are best tested by examining changes to average conditions that are large relative to the seasonal signal. In addition, well-understood changes in solar heating (Milankovitch cycles) occur on long timescales, providing an excellent natural experiment to examine climate responses to perturbations in the Earth's radiative balance.

Relative to today, the Pliocene warm period was characterized by: ~3 °C higher global surface temperatures, 10–20 m higher sea level, enhanced thermohaline circulation^{1,2}, slightly reduced Antarctic ice sheets, emerging but small Northern Hemisphere ice coverage³, and slightly (30%) higher atmospheric carbon dioxide concentrations^{1,4}. A small decrease in carbon dioxide concentration could explain the cooling at the end of the warm period if coupled with positive feedbacks, as suggested for the onset of significant Antarctic glaciation⁵. However, whether these feedbacks primarily involved low- or high-latitude processes has been controversial.

Although high-latitude feedbacks (for example, related to ocean heat transport or ice albedo^{6,7}) may have accelerated cooling once NHG began, the impact of glaciation on global-scale cooling still needs to be explored. Alternatively, long-term reorganization of tropical conditions could have strongly influenced global climate, as occurs interannually with the El Niño–Southern Oscillation phenomenon^{8,9}. Even small changes in tropical temperature patterns can profoundly affect extratropical conditions on geological timescales¹⁰. Thus, low-latitude tectonic events (restriction of Panamanian or Indonesian seaways) may have changed the distribution of heat between basins^{7,11,12}, causing reorganization of climate patterns, the end of the warm period, and ultimately

intensification of NHG. Yet the global impact of these tectonic events has not been adequately examined. Finally, bidirectional high–low latitude interactions may explain important features of the transition. For example, intensification of NHG could have resulted in cooler deepwaters (formed at high latitudes) and a subsequent increase in deep ocean stratification. Increased stratification may have caused the ventilated thermocline to shoal, allowing cold water to upwell in tropical and subtropical regions, thereby altering global climate patterns⁸.

To test hypotheses that explain the end of the warm period, we compare distant palaeoceanographic records to examine tropical–extratropical interactions. This analysis results in the fundamental conclusion that major long-term cooling steps in different regions (for example, intensification of NHG, reorganization of tropical circulation) did not all occur at the same time. Thus, regionally specific processes caused cooling phases at different times, and the end of the warm period was not forced by a single episodic event whose effects propagated globally^{7,11}. Rather, it must have been forced gradually. This conclusion is backed up by the analysis of changes in the amplitude of high-latitude climate variability, which indicates that climate sensitivity also increased gradually.

High-latitude climate trends

The oxygen isotopic ($\delta^{18}\text{O}$) composition of benthic foraminifera reflects deepwater temperature and ice volume changes, both indicative of high-latitude climate change. The $\delta^{18}\text{O}$ record (Fig. 1a) indicates that during the Pliocene warm period, high-latitude climate was, on average, warmer than interglacial periods of the past 1 Myr. The onset of significant NHG, reflected by the first obliquity-related cycles with high $\delta^{18}\text{O}$ values, or the beginning of the '41-kyr world', occurred ~2.75 Myr ago (Fig. 1a). These intense glaciations evident in the $\delta^{18}\text{O}$ record caused pronounced ice-sheet calving as documented by North Pacific¹³ (Fig. 2b) and North Atlantic¹⁴ ice-rafted debris records. Other evidence of major Northern Hemisphere high-latitude climate reorganization includes the sudden increase in North Pacific surface-water stratification¹³

(Fig. 2c), and pronounced reductions in North Atlantic^{1,15} and Pacific¹⁶ thermohaline overturn rates.

Subtropical climate trends

Two major steps of climate reorganization are evident in many records of subtropical cooling at the end of the Pliocene warm period. Both cooling steps are well expressed in the best-studied subtropical continental regions: those influenced by the Asian and African (Fig. 2d) monsoon systems^{17–19}. The first step (between 3.0 and 2.5 Myr ago) was coincident with the onset of significant NHG. However, the second step (between 2.0 and 1.5 Myr ago) occurred well after the onset of significant of NHG. Sea surface temperature (SST) in the West African upwelling system also began to decrease around 3.0 Myr ago (ref. 20) with the majority of the 10 °C-cooling trend after the onset of significant NHG (Fig. 2e). In all, these records indicate that at least the first step in subtropical climate change may be related to NHG. However, the second step in subtropical climate change occurred when high-latitude climate was relatively stationary (Fig. 1a).

To characterize further the subtropical climate change in the Pliocene, we generated new California margin records (Fig. 2f–i). Ocean Drilling Program (ODP) Site 1014 (32°50'N, 119°59'W, 1165 m water depth) is located in a sensitive region between the warm North Pacific subtropical gyre and the highly productive cooler upwelling conditions of the California margin. Our records of calcite mass accumulation rate (CaCO₃-MAR) and a seasonality proxy (see Methods) are used to monitor the evolution of the upwelling system, and like records from other subtropical regions, indicate that changes occurred in two steps over the last 4 Myr. The first step, an increase in CaCO₃-MAR just after 3.0 Myr ago, was not accompanied by a change in seasonality. The second step, a pronounced decrease in CaCO₃-MAR and increase in seasonality, occurred about 1.7 Myr ago. The first-order trends in CaCO₃-MAR are unlikely to be an artefact of changes in dissolution²¹ given the shallow water depth and unchanging bottom water conditions of this site¹⁶ (and preliminary measurements of alkenone-MAR, an organic marker of coccolithophore productivity, are in agreement with the CaCO₃-MAR record). Nor are the trends related to errors

in estimating sedimentation rate, because they are reproducible using independent dating methods (for example, high-resolution correlations, biostratigraphy), and because the change 1.7 Myr ago is also evident in records (%CaCO₃, %C_{org}, seasonality) (Fig. 2f, g, h) that do not depend on accurate estimates of sedimentation rate.

Today, CaCO₃-MAR on the California coast is greatest in offshore

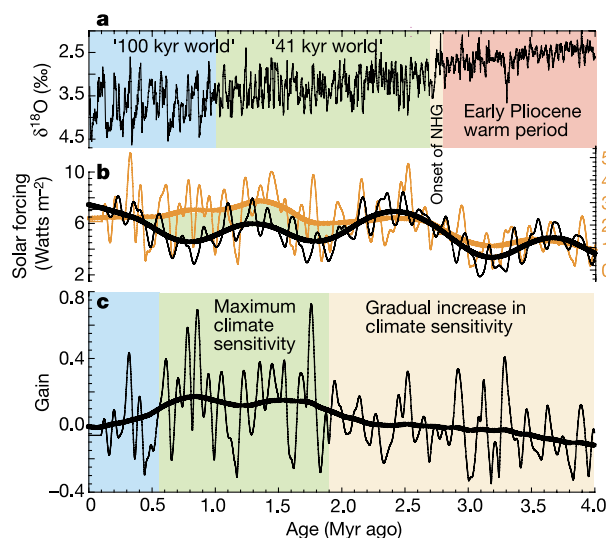


Figure 1 Evolution of high-latitude climate and sensitivity (adapted from ref. 44). **a**, Benthic foraminifera $\delta^{18}\text{O}$ record of high-latitude climate change. **b**, Deconvolution of the obliquity component of scaled oxygen isotope (orange curve) and of solar forcing (black curve) records (smoothed records are thick lines). The period of greatest high-latitude response to solar forcing is shaded in green. **c**, Gain record, a measure of climate sensitivity, is the ratio of two deconvoluted records. Note that the transition to the high-amplitude $\delta^{18}\text{O}$ cycles of the '41-kyr world' was not related to a change in sensitivity.

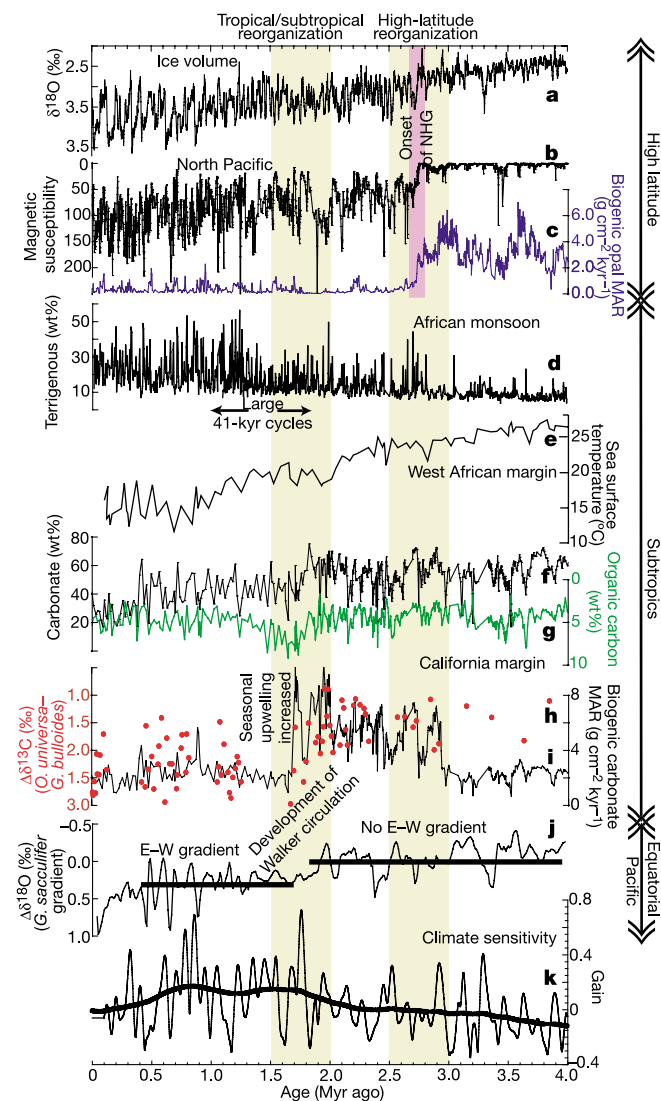


Figure 2 Comparison of climate evolution in different regions. **a**, Benthic foraminifera $\delta^{18}\text{O}$ record from Fig. 1. **b**, Magnetic susceptibility record from Subarctic Pacific ODP Site 882 (50° N, 168° E; water depth 3,244 m) reflects ice-rafted debris input to sediments¹³. **c**, Mass accumulation rate (MAR) of biogenic opal from ODP Site 882 interpreted as inversely related to surface stratification¹³. **d**, Terrigenous wt% at ODP Site 664¹⁸ (6° N, 23° W; water depth 3,806 m) indicates increased aridity at the onset of NHG and further aridification just after ~2.0 Myr ago. Time interval when 41-kyr cycles dominate variability is annotated. **e**, Alkenone-saturation-index-derived SST record from ODP Site 1084 (26° S, 13° E; water depth 1,992 m) reflects strength of West African margin upwelling²⁰. **f**, Carbonate (wt%) from ODP Site 1014. **g**, Organic carbon (wt%) from ODP Site 1014. **h**, Carbon isotopic difference between *O. universa* and *G. bulloides*, two species of planktonic foraminifera that proliferate in different seasons from ODP Site 1014, reflects surface water seasonality (red dots). **i**, MAR of biogenic carbonate from ODP Site 1014 reflects the accumulation rate of carbonate in the sediments. **j**, Oxygen isotope difference (measured on *G. sacculifer*) between eastern Pacific ODP Site 851 (ref. 26) (3° N, 111° W, water depth 3,761 m) and western Pacific ODP Site 806 (refs 49, 50) (0°, 159° E, water depth 2,521 m) reflects the development of a hydrographic gradient across the tropical Pacific. **k**, Gain record, a measure of climate sensitivity (from Fig. 1).

windstress-curl-driven upwelling regions, where well-stratified surface waters²² and coccolithophore productivity develop seasonally. Coccolithophore productivity is favoured in the subsurface chlorophyll maximum, where cool nutrient-rich waters from the underlying ventilated thermocline (the strong thermal gradient between warm surface waters and cool deep water) upwell into the photic zone. Thus, the increase in CaCO_3 -MAR approximately coincident with onset of significant NHG may reflect shoaling of the ventilated thermocline or an increase in windstress curl at the edge of the North Pacific gyre, without a notable change in surface seasonality. These conditions allowed for the delivery of subsurface nutrient-rich waters into the photic zone to support coccolithophore production, while maintaining well-stratified conditions. 1.7 Myr ago, a further increase in wind-driven upwelling or shoaling of the thermocline, accompanied by decreased stratification at least seasonally, as occurs today, would have caused total CaCO_3 production to drop and seasonality to increase. This second major change on the California margin and in other subtropical regions occurred when long-term high-latitude climate was relatively stable (Fig. 2a–c).

Tropical climate trends

The state of the tropics is intimately tied to subsurface thermocline conditions. The source of thermocline waters is surface water subducted at mid-latitudes²³. Today, the thermocline is sufficiently shallow and cool at the eastern boundaries of the tropical and subtropical Pacific and Atlantic Oceans that upwelling results in cool SSTs in these regions, setting up west–east gradients in SST and pressure. In the tropical Pacific, the west–east gradients strengthen the easterly trade winds and reinforce cool upwelling in the east, thereby further augmenting the temperature and pressure gradients. This positive air–sea feedback is required to maintain strong zonal, or Walker, circulation. Small perturbations that weaken Walker circulation are amplified by these same air–sea feedbacks, causing the thermocline to deepen in the east and El Niño conditions to develop²⁴. Teleconnections then cause climate anomalies in the tropical regions to influence climate patterns on a global scale.

Although limited data compilations indicate that the tropics were probably not, on average, significantly warmer than today²⁵ during the early Pliocene, there is evidence that the pattern of tropical conditions resembled a permanent El Niño: the thermocline was deep, and the west–east SST gradient was greatly reduced compared with modern normal conditions^{26–28}. Furthermore, long-term extratropical climate patterns in the early Pliocene were similar to those manifested during a modern El Niño¹². Mid-latitude surface waters, which are subducted into the thermocline thereby controlling its character^{23,29}, were several degrees warmer than today in the Pliocene warm period²⁶. With the thermocline either deeper or warmer than today, wind-driven upwelling would not have cooled the eastern tropical Pacific sufficiently to maintain strong Walker circulation. Thus, mid-latitude warmth, and tropical–extratropical coupling via the thermocline, could explain the weak Walker circulation of the Pliocene warm period.

The possible effects of tropical processes on extratropical climate events (for example, NHG) can be assessed by considering the timing of tropical climate changes and the development of strong Walker circulation. Tropical records indicate that significant tropical climate reorganization occurred twice: once between 4.5 and 4.0 Myr ago, well before significant NHG, and once between 2.0 and 1.5 Myr ago, well after the onset of significant NHG. Between 4.5 and 4.0 Myr ago, a marked shift in surface-water hydrographic gradients between the Pacific and Caribbean³⁰ and between the Caribbean and the western tropical Atlantic³¹, shoaling of the thermocline in the east Pacific²⁸, and circulation changes in the Atlantic^{32–34}, were possibly forced by tectonic event(s) (for example, restriction of the Panamanian³⁰ or Indonesian¹¹ seaways). However, the SST gradient across the Pacific (Fig. 2i) did not increase until between 2.0 and 1.5 Myr ago^{26,27}, indicating that the

development of strong Walker circulation was potentially related to the second major transition in subtropical conditions (Fig. 2d, e, h, i).

Gradual global cooling

In the following section, we integrate important aspects of palaeoclimate observations from around the globe, and conclude that the transition from the Pliocene warm period to the Pleistocene cold period cannot be explained by a single tectonic or threshold event, but rather, was forced gradually. The onset of significant NHG (~2.75 Myr ago) (including major changes in thermohaline circulation and deep ocean density and nutrient stratification^{1,2,15,16}), occurred at the same time as subtropical regions began to cool, confirming a strong linkage between high latitude and subtropical regions^{17–20}. However, the onset of significant NHG (Fig. 2b, c) was not related to fundamental changes in tropical conditions. It occurred when tropical conditions were stable and El Niño-like, about a million years after an initial, possibly tectonically driven change due to a closing tropical seaway, and about a million years before the establishment of strong Walker circulation (Fig. 2j). Thus, the rapid transition ~2.75 Myr ago that characterizes climate evolution in high latitude and some subtropical regions, must have strictly involved extratropical processes such as those involving ice albedo, thermohaline circulation, monsoon strength, or biogeochemical¹³ changes, and was potentially initiated by reaching a greenhouse gas^{35,36}, ice-sheet substrate³⁷, or tectonic³⁸ threshold. A different threshold was reached just after 2.0 Myr ago that switched the tropics and subtropics (Fig. 2i) into the modern mode of circulation with relatively strong Walker circulation and cool subtropical temperatures.

Although the asynchronicity of climate change steps attests to the importance of regionally distinct processes to explain those steps, they do not preclude the existence of important feedbacks between distant regions. For example, the observed increase in wind-driven subtropical upwelling could be related to the increase in the latitudinal temperature gradient³⁹ that occurred first with high-latitude cooling 2.75 Myr ago, and then with the enhancement of the western tropical warm pool just after 2.0 Myr ago. However, recent modelling results⁴⁰ predict that high-latitude cooling should have caused the subtropical winds to weaken. This apparent contradiction could indicate that subtropical upwelling records reflect changes either in the position, rather than the strength, of the subtropical gyre or in thermocline conditions, rather than changes in wind strength. Thus, the ocean's thermocline, rather than the atmosphere, may have served as a link between distant locations. A recent study emphasizes that global cooling of the past 3 Myr was related to changes in the depth of the ventilated thermocline⁸. The study asserts that as high-latitude deepwater formation regions cooled, deepwater temperature decreased, ocean stratification increased, and the thermocline shoaled. This caused tropical and subtropical upwelling regions to cool, invigorating west–east temperature and pressure gradients and reinforcing the cooling trend. The observed changes in Pacific subtropical and tropical upwelling regions ~2.0 Myr ago suggests that distant locations are being influenced by a shoaling thermocline and supports the hypothesis. The hypothesis is also supported by evidence that during the warm period relative to today, deep ocean temperature was high⁴¹ and ocean stratification was reduced¹⁶ at the same time that the thermocline was deep²⁶. However, further testing of the idea using the climate change of the past 4 Myr awaits the generation of more data on the detailed history and global nature of changes in deep ocean stratification, the depth of the ventilated thermocline, and the temperature of upwelling regions.

Although the changes in the ventilated thermocline⁸ and the temperature gradient³⁹ each describe possible linkages between distant locations and positive feedbacks to cooling, they provide no explanation of the ultimate cause of global cooling. Certainly, the fact that the global cooling transition occurred with regionally

distinct timing suggests that it was not caused by a single event. A single event could not have acted as a 'trigger' that initiated the transition, as it is unlikely that subsequent interactions between climate system components over a million years or so could have by themselves 'fortuitously' led to global cooling. Thus, gradual changes in atmospheric greenhouse gas composition, basin geometry, or land-surface conditions, occurring at least over the past ~4–1.5 Myr, most probably forced global cooling and need further investigation.

Climate sensitivity

Glacial–interglacial cycles have been a striking feature of climate change for the past 4 Myr (Fig. 1a). What controls the frequency of these cycles? Variability of seasonal solar heating at all latitudes is predominantly controlled by precession of the Earth's position on its orbit during summer, with cycles of 19–23 kyr and amplitudes often over 100 W m^{-2} . However, only 12% of the variance in the $\delta^{18}\text{O}$ record, before 1 Myr ago, is in the precession band, whereas 43% is in the obliquity band (cycles of ~41 kyr). Thus, variations in solar heating due to changes in the Earth's angle of tilt, or obliquity, must be an important driving force of glacial cycles before the Late Pleistocene period^{8,42}. In fact, although precession controls seasonal heating, obliquity controls over 85% of the variance in the annual average solar heating at all latitudes and in the high-to-low latitudinal gradient of solar heating. Clearly, the frequency of glacial cycles was in large part controlled by obliquity-related solar variations; but why did the amplitude of glacial cycles change through time? Comparing the non-stationary $\delta^{18}\text{O}$ record to the long-term regional trends of the past 4 Myr can address this question.

Because the onset of significant NHG and increased amplitude of obliquity-related $\delta^{18}\text{O}$ variations (the beginning of the '41-kyr world') both occurred ~2.75 Myr ago (Fig. 1a), it seems that there was a causal relationship between average high-latitude conditions and the amplitude of glacial–interglacial cycles. However, an evaluation of changes in the climate's response to solar forcing must account for how the forcing itself changed. In fact, the amplitude of obliquity solar cycles was not constant; it was modulated by interactions with eccentricity cycles of the Earth's orbit⁴³. The amplitude modulation of obliquity cycles in both the forcing (solar) and response ($\delta^{18}\text{O}$) records was quantified using complex demodulation performed on the filtered 41-kyr components of these records⁴⁴ (Fig. 1b). The gain (the ratio of the two demodulated records) (Fig. 1c) reflects the amplitude of the $\delta^{18}\text{O}$ response relative to solar forcing, thereby providing a measure of high-latitude climate sensitivity. The increase in the amplitude of $\delta^{18}\text{O}$ variability between 3.0 and 2.5 Myr ago (Fig. 1b), or the beginning of the '41-kyr world', was a direct response to the increasing amplitude of solar forcing, and therefore cannot be directly attributed to the onset of significant NHG. Climate sensitivity increased gradually after 4.0 Myr ago, culminating in a period of highest sensitivity after ~2.0 Myr ago. A calculation of the sensitivity, or gain, using the precessional components of the $\delta^{18}\text{O}$ and solar-forcing records provides the same result.

Comparison of the sensitivity record (Fig. 2k) to changes in the mean state of climate in different regions leads us to several conclusions. Sensitivity gradually increased and then reached its highest level during the '41-kyr world' (Figs 1, 2k), even while average high-latitude conditions remained relatively stable (Fig. 2a–c). Thus, the strength of feedbacks that increasingly amplified solar cycles after 4.0 Myr ago was probably independent of long-term high-latitude conditions such as ice-sheet size or deepwater formation strength. In contrast, the approximate temporal correlation between the enhancement of Walker circulation, seasonal subtropical upwelling, and climate sensitivity, just after ~2.0 Myr ago, suggests that important processes that amplify obliquity-related solar forcing may reside in tropical and/or subtropical regions. As upwelling regions cooled, potent air–sea feedbacks associated with

the maintenance of Walker circulation and subtropical land–sea pressure gradients, amplified small perturbations to the radiative forcing resulting from changes in Earth's tilt. This idea is consistent with the low-amplitude tropical climate cycles, and their lack of coherency to solar forcing, before the establishment of Walker circulation²⁷. It is also consistent with the observation that obliquity-related cycles dominated African monsoon variability after 1.8 Myr ago¹⁹ (Fig. 2d). To test our conclusions, a quantitative assessment of changes in low-latitude climate sensitivity to solar forcing is needed once long high-resolution records from tropical and subtropical regions are generated. Furthermore, since about 35% of the variance in the $\delta^{18}\text{O}$ record before 1 Myr ago is not directly related to orbital solar forcing, a thorough evaluation of 'sensitivity' should also focus on the source of higher-frequency variations.

The observation that average lower-latitude conditions influenced high-latitude climate sensitivity has implications on hypotheses that explain how obliquity solar cycles might drive climate change. One hypothesis⁸, generally supported by our observations, predicts that thermocline temperature was influenced by annual average heating (controlled by obliquity) at mid-latitudes where thermocline waters are subducted. Changes in thermocline temperature influenced SST in tropical upwelling regions, which influenced high-latitude climate through teleconnections. The climate response to solar forcing was weak in the Pliocene warm period because the thermocline was too deep to influence SST in upwelling regions, regardless of obliquity-forced perturbations. However, the sensitivity (Fig. 2k) increased as the long-term average conditions of the thermocline shoaled or cooled gradually. A different theory⁴² suggests that low-to-high latitude gradient in solar forcing, coupled with strong ice-albedo feedbacks, controlled obliquity-related climate change. However, our data indicate that if the gradient in solar forcing controlled glacial cycles in the late Pliocene, it may have instead done so with the help of tropical or subtropical processes.

Implications for understanding climate change

Several lessons can be drawn from the comparison of Plio–Pleistocene climate change records from distant locations. First, although changes in forcing were gradual, strong regional nonlinear responses generated pronounced regional climate changes including the onset of significant NHG. Second, the ventilated thermocline and/or latitudinal temperature gradient may have played an important role in linking subtropical conditions to change in other regions. Finally, tropical and subtropical conditions, specifically the time-averaged strength of coldwater upwelling in the eastern Pacific, and of Walker circulation, had a strong influence on the climate response to radiative changes. Thus, the last 4 Myr illustrates that as globally average conditions change, so do the feedbacks or 'rules' that determine climate sensitivity. This conclusion is relevant to studies of future global warming because it emphasizes the importance of 'background' or average tropical conditions in predicting high-frequency climate change. Furthermore, understanding processes responsible for recent climate change of the last hundreds or thousands of years, when average background conditions changed very little, is unlikely to be sufficient to predict climate variability for periods with different globally averaged conditions. This highlights the importance of developing theory to explain ocean and atmospheric change, and testing that theory using records from geologic time periods that represent a large dynamic range of climate conditions. □

Methods

Measurements of $\delta^{18}\text{O}$ on benthic foraminifera (*Cibicides mckennai*) were made in the stable isotope facility at the University of California, Santa Cruz on samples from ODP Site 1014, with an approximate average resolution of ~3–4 kyr (data not shown). The age of the sediment, from 0–3.0 Myr ago, was determined by correlating this new $\delta^{18}\text{O}$ record to

the $\delta^{18}\text{O}$ record from ODP Site 846 (ref. 45). Because there were core gaps in the sedimentary section older than 1.0 Myr ago, the age model was further refined by tuning magnetic susceptibility⁴⁶ (measured with the down-hole Geologic High Resolution Magnetic Tool on Hole 1014A after drilling) to obliquity⁴³. Measurements of weight per cent of calcium carbonate (%CaCO₃) (Fig. 2f) and organic carbon (%C_{org}) (Fig. 2g) of ODP Site 1014 samples were made at Boise State University using a UIC coulometer attached to a modified version of the CM-5120 furnace module, as described in a previous study⁴⁷. The CaCO₃ component at this site is made primarily of coccoliths⁴⁶. The 'seasonality' (Fig. 2h) is the mathematical difference between carbon isotopic values ($\Delta\delta^{13}\text{C}$) measured on two species of planktonic foraminifera, *Orbulina universa* and *Globigerina bulloides*, that favour different seasonal conditions⁴⁸. The record of CaCO₃-MAR (Fig. 2i) was calculated using %CaCO₃, sedimentation rate, and ODP Site 1014 shipboard dry bulk density measurements⁴⁶. All ODP Site 1014 data will be archived on the NDCG-NOAA palaeoclimate database website: www.ngdc.noaa.gov/paleo/paleo.html.

Received 6 November 2003; accepted 7 April 2004; doi:10.1038/nature02567.

1. Raymo, M. E., Grant, B., Horowitz, M. & Rau, G. H. Mid-Pliocene warmth: Stronger greenhouse and stronger conveyor. *Mar. Micropaleo.* **27**, 313–326 (1996).
2. Ravelo, A. C. & Andreasen, D. H. Enhanced circulation during a warm period. *Geophys. Res. Lett.* **27**, 1001–1004 (2000).
3. Haywood, A. M., Valdes, P. J. & Sellwood, B. W. Global scale palaeoclimate reconstruction of the middle Pliocene climate using the UKMO GCM: initial results. *Glob. Planet. Change* **25**, 239–256 (2000).
4. Van der Burgh, J., Visscher, H., Dilcher, D. L. & Kürschner, W. M. Paleatmospheric signatures in Neogene fossil leaves. *Science* **260**, 1788–1790 (1993).
5. DeConto, R. M. & Pollard, D. Rapid Cenozoic glaciation of Antarctica induced by declining atmospheric CO₂. *Nature* **421**, 245–249 (2003).
6. Raymo, M. E., Rind, D. & Ruddiman, W. F. Climatic effects of reduced Arctic sea ice limits in the GISS II general circulation model. *Paleoceanography* **5**, 367–382 (1990).
7. Driscoll, N. W. & Haug, G. H. A short circuit in thermohaline circulation; a cause for Northern Hemisphere glaciation? *Science* **282**, 436–438 (1998).
8. Cane, M. A. & Evans, M. Do the tropics rule? *Science* **290**, 1107–1108 (2000).
9. Philander, S. G. & Fedorov, A. V. Role of tropics in changing the response to Milankovitch forcing some three million years ago. *Paleoceanography* **18**, 23–1–23–11 (2003).
10. Yin, J. H. & Battisti, D. S. The importance of tropical sea surface temperature patterns in simulations of last glacial maximum climate. *J. Clim.* **14**, 565–581 (2001).
11. Cane, M. A. & Molnar, P. Closing of the Indonesian Seaway as a precursor to East African aridification around 3–4 million years ago. *Nature* **411**, 157–162 (2001).
12. Molnar, P. & Cane, M. A. El Niño's tropical climate and teleconnections as a blueprint for pre-Ice Age climates. *Paleoceanography* **17**, 11–11 (2002).
13. Haug, G. H., Sigman, D. M., Tiedemann, R., Pedersen, T. F. & Sarnthein, M. Onset of permanent stratification in the subarctic Pacific Ocean. *Nature* **401**, 779–782 (1999).
14. Ruddiman, W. F. & Raymo, M. E. Northern Hemisphere climate regimes during the past 3 Ma: possible tectonic connections. *Phil. Trans. R. Soc. Lond. B* **318**, 411–430 (1988).
15. Raymo, M. E., Hodell, D. & Jansen, E. Response of deep ocean circulation to initiation of northern hemisphere circulation (3–2 Ma). *Paleoceanography* **7**, 645–672 (1992).
16. Kwiek, P. & Ravelo, A. C. Pacific Ocean intermediate and deep water circulation during the Pliocene. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **154**, 191–217 (1999).
17. Clemens, S., Murray, D. W. & Prell, W. L. Nonstationary phase of the Plio-Pleistocene Asian monsoon. *Science* **274**, 943–948 (1996).
18. DeMenocal, P. B. Plio-Pleistocene african climate. *Science* **270**, 53–59 (1995).
19. Tiedemann, R., Sarnthein, M. & Shackleton, N. J. Astronomic timescale for the Pliocene Atlantic $\delta^{18}\text{O}$ and dust flux records of Ocean Drilling Program site 659. *Paleoceanography* **9**, 619–638 (1994).
20. Marlow, J. R., Lange, C. B., Wefer, G. & Rosell-Melé, A. Upwelling intensification as part of the Plio-Pleistocene climate transition. *Science* **290**, 2288–2294 (2000).
21. Ravelo, A. C. *et al.* Pliocene carbonate accumulation along the California Margin. *Paleoceanography* **12**, 729–741 (1997).
22. Okada, H. & Honjo, S. The distribution of oceanic coccolithophorids in the Pacific. *Deep-Sea Res.* **20**, 355–374 (1973).
23. Harper, S. Thermocline ventilation and pathways of tropical-subtropical water mass exchange. *Tellus A* **52**, 330–345 (2000).
24. Bjerknes, J. Atmospheric teleconnections from the tropical Pacific. *Mon. Weath. Rev.* **97**, 163–172 (1969).
25. Dowsett, H., Barron, J. & Poore, R. Z. Middle Pliocene sea surface temperatures: A global reconstruction. *Mar. Micropaleo.* **27**, 13–25 (1996).
26. Cannariato, K. & Ravelo, A. C. Plio-Pleistocene evolution of eastern tropical Pacific surface water

- circulation and thermocline depth. *Paleoceanography* **12**, 805–820 (1997).
27. Wara, M. *On the Interaction Between Tropical and High Latitude Climate Over the Past 5 Million Years*. Thesis, Univ. California, Santa Cruz (2003).
28. Chaisson, W. & Ravelo, A. C. Pliocene development of the East-West hydrographic gradient in the Equatorial Pacific. *Paleoceanography* **15**, 497–505 (2000).
29. Gu, D. & Philander, S. Interdecadal climate fluctuations that depend on exchanges between the tropics and extratropics. *Science* **275**, 805–807 (1997).
30. Haug, G. H. *et al.* Role of Panama uplift on oceanic freshwater balance. *Geology* **29**, 207–210 (2001).
31. Billups, K. *et al.* A link between ocean heat transport, thermohaline circulation and Intertropical Convergence Zone in the early Pliocene Atlantic. *Geology* **28**, 319–322 (1999).
32. Chaisson, W. & Ravelo, A. C. Changes in upper water-column structure at Site 925, Late Miocene–Pleistocene: Planktonic foraminifer assemblage and isotopic evidence. *Proc. Ocean Drill. Prog. Sci. Res.* **154**, 255–268 (1997).
33. Norris, R. D. Miocene–Pliocene surface-water hydrography of the eastern equatorial Atlantic. *Proc. Ocean Drill. Prog. Sci. Res.* **159**, 539–555 (1998).
34. Billups, K., Ravelo, A. C. & Zachos, J. C. Early Pliocene climate: A perspective from the western equatorial Atlantic warm pool. *Paleoceanography* **13**, 459–470 (1998).
35. Saltzman, B. & Verbitsky, M. Y. Multiple instabilities and modes of glacial rhythmicity in the Plio-Pleistocene: a general theory of late Cenozoic climatic change. *Clim. Dyn.* **9**, 1–15 (1993).
36. Raymo, M. E. Glacial puzzles. *Science* **281**, 1467–1468 (1998).
37. Clark, P. U., Alley, R. B. & Pollard, D. Northern hemisphere ice-sheet influences on global climate change. *Science* **286**, 1104–1111 (1999).
38. Ruddiman, W. F. & Raymo, M. E. Northern hemisphere climate régimes during the past 3 Ma: possible tectonic connections. *Phil. Trans. R. Soc. Lond. B* **318**, 411–430 (1988).
39. Rind, D. Relating paleoclimate data and past temperature gradients: Some suggestive rules. *Quat. Sci. Rev.* **19**, 381–390 (2000).
40. Haywood, A. M., Valdes, P. J. & Sellwood, B. W. Global scale palaeoclimate reconstruction of the middle Pliocene climate using the UKMO GCM: initial results. *Glob. Planet. Change* **25**, 239–256 (2000).
41. Lear, C. H., Rosenthal, Y. & Wright, J. D. The closing of a seaway: ocean water masses and global climate change. *Earth Planet. Sci. Lett.* **210**, 425–436 (2003).
42. Raymo, M. E. & Nisancioglu, K. The 41 kyr world: Milankovitch's other unsolved mystery. *Paleoceanography* **18**, 1011 (2003).
43. Laskar, J., Joutel, F. & Boudin, F. Orbital, precessional, and insolation quantities for the Earth from –20 Myr to +10 Myr. *Astron. Astrophys.* **270**, 522–533 (1993).
44. Andreasen, D. H. *Isolating Sources of Plio-Pleistocene Paleoclimate Variability*. Thesis, Univ. California, Santa Cruz (2001).
45. Shackleton, N. J., Crowhurst, S., Hageberg, T., Pisias, N. G. & Schneider, D. A. A new Late Neogene time scale: application to Leg 138 Sites. *Proc. Ocean Drill. Prog. Sci. Res.* **138**, 73–101 (1995).
46. Lyle, M. *et al.* Site 1014. *Proc. Ocean Drill. Prog. Init. Rep.* **167**, 175–222 (1997).
47. Lyle, M. *et al.* Millennial-scale CaCO₃ and C_{org} events along the northern and central California margin: Stratigraphy and origins. *Proc. Ocean Drill. Prog. Sci. Res.* **167**, 163–182 (2000).
48. Sautter, L. R. & Thunell, R. C. Seasonal variability in the $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ of planktonic foraminifera from an upwelling environment: sediment trap results from the San Pedro Basin, Southern California Bight. *Paleoceanography* **6**, 307–334 (1991).
49. Berger, W. H., Bickert, T., Schmidt, H. & Wefer, G. Quaternary oxygen isotope record of pelagic foraminifera; Site 806, Ontong Java Plateau. *Proc. Ocean Drill. Prog. Sci. Res.* **130**, 381–395 (1993).
50. Jansen, E., Mayer, L. A., Backman, J., Leckie, R. M. & Takayama, T. Evolution of Pliocene climate cyclicity at Hole 806B (5–2 Ma); oxygen isotope record. *Proc. Ocean Drill. Prog. Sci. Res.* **130**, 349–362 (1993).

Acknowledgements We thank L. Anderson, M. Cane, B. Chaisson, P. Dekens, G. Haug, K. Lawrence, P. Molnar and G. Philander, for suggestions that influenced this work. M. Flower, Z. Irons and K. Ziegler helped with the data generation. G. Haug and P. deMenocal supplied their data for Fig. 2. The National Science Foundation funded this work. Samples were provided by the Ocean Drilling Program.

Authors' contributions D.H.A. conceived of the ideas presented in Fig. 1. M.L. and A.O.L. were responsible for carbonate and organic carbon analyses. All authors contributed scientific ideas. A.C.R. was primarily responsible for generating isotope data on ODP Site 1014 and constructing age model, formulating the main conclusions, and writing the paper.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.C.R. (acr@es.ucsc.edu).

Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes

Jeffrey S. Han¹, Suzanne T. Szak² & Jef D. Boeke¹

¹Department of Molecular Biology and Genetics and High Throughput Biology Center, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

²Biogen, Inc., Cambridge, Massachusetts 02142, USA

LINE-1 (L1) elements are the most abundant autonomous retrotransposons in the human genome, accounting for about 17% of human DNA. The L1 retrotransposon encodes two proteins, open reading frame (ORF)1 and the ORF2 endonuclease/reverse transcriptase. L1 RNA and ORF2 protein are difficult to detect in mammalian cells, even in the context of overexpression systems. Here we show that inserting L1 sequences on a transcript significantly decreases RNA expression and therefore protein expression. This decreased RNA concentration does not result from major effects on the transcription initiation rate or RNA stability. Rather, the poor L1 expression is primarily due to inadequate transcriptional elongation. Because L1 is an abundant and broadly distributed mobile element, the inhibition of transcriptional elongation by L1 might profoundly affect expression of endogenous human genes. We propose a model in which L1 affects gene expression genome-wide by acting as a 'molecular rheostat' of target genes. Bioinformatic data are consistent with the hypothesis that L1 can serve as an evolutionary fine-tuner of the human transcriptome.

L1 elements are the most abundant autonomous retrotransposons in the human genome, accounting for about 17% of human DNA¹. A functional full-length L1 element, depicted in Fig. 1a, contains an internal promoter in the 5' untranslated region (5' UTR) that initiates transcription at base 1 (ref. 2). This is followed by two long open reading frames, ORF1 and ORF2, required for retrotransposition^{3,4} *in cis*^{5,6}. ORF1 encodes an RNA-binding protein that has nucleic acid chaperone activity *in vitro*⁷⁻⁹, but no known specific role in the L1 replication mechanism. ORF2 encodes a protein with endonuclease and reverse transcriptase activities, both critical for retrotransposition^{3,4}. A short 3' UTR is followed immediately by a poly(A) tail, and the entire element is typically flanked by target site duplications. The L1 retrotransposition machinery is not only used for L1 mobilization itself but also assists in the retrotransposition of Alu retroelements¹⁰.

L1 elements are mobilized through target-primed reverse transcription, in which ORF2 nicks target DNA, using the resultant 3'-OH to prime the reverse transcription of L1 RNA^{11,12}. ORF2 is probably relatively non-processive and often fails to reach the RNA 5' end during first-strand synthesis. This explains, at least in part, the distribution of predominantly non-functional, 5' truncated L1s in the genome^{13,14}.

A puzzling feature of L1 is the difficulty in detecting L1 RNA and ORF2 protein in mammalian cells, even in the context of high-copy plasmids and overexpression. In contrast, recombinant ORF2 was expressed successfully in yeast and baculovirus systems^{12,15,16}. This suggests a mammalian-specific mechanism for negatively regulating L1 expression, which, considering the mutagenic nature of transposition and dearth of L1 expression in somatic cells^{17,18}, is not unexpected. The difficulty in expressing L1 RNA suggests a transcriptional defect, but translation might also have a function in low ORF2 expression, because the ORF2 translation initiation mechanism is not understood. Recent data indicate that L1 transcription might be inhibited by cryptic premature polyadenylation signals¹⁹, which would produce nonfunctional truncated L1 transcripts. Here we show that poor expression of ORF2 results from the inability of RNA polymerase to elongate efficiently through L1 coding

sequences, with polyadenylation having a smaller role. We also show that L1 sequences in the antisense orientation inhibit transcription by producing premature polyadenylation; this is significant because most L1 elements within human genes are in the antisense orientation.

L1 forms a component of most mammalian transcription units, but the effects of these primarily intronic inserted sequences have not been studied carefully. About 79% of human genes are estimated to contain at least one segment of L1 in their transcription unit (see Methods), and L1 segments from pre-existing and newly derived insertions usually contain L1 ORF2 (refs 14, 20–22). As these sequences are mostly intronic, it has been assumed that the extra sequences are spliced out and do not affect target gene expression. Our findings indicate that, through a combination of transcriptional elongation inhibition and premature polyadenylation, L1 insertions in either orientation can affect the RNA production of endogenous genes, both qualitatively and quantitatively. We propose a human genome model in which L1 has led to numerous subtle but potentially significant transcriptome alterations.

ORF2 reduces RNA and protein amounts

To examine the effect of ORF2 sequences on expression, we fused either *lacZ* or L1 ORF2 coding regions downstream of the green fluorescent protein (GFP) ORF (Fig. 1b). Placing these test transcripts downstream of GFP permitted the examination of RNA and protein expression independently of translation initiation, because translation initiates in GFP. The presence of ORF2 sequence significantly lowered steady-state protein production and led to a 70-fold decrease in RNA relative to *lacZ* (Fig. 1c, d, lanes 2 and 3). Inserting ORF2 in the antisense orientation produced a similar, but less potent, decrease in full-length RNA. These phenomena also occur with the mouse ORF2 sequence (Fig. 1d, lanes 5 and 6) and are therefore not human L1-specific. The effect is limited neither to specific cell lines nor to the promoter used (Supplementary Fig. S2a, b).

When L1 ORF2 is inserted in the antisense orientation, most

of the decrease in full-length RNA is accounted for by lower-molecular-mass species. These are probably prematurely polyadenylated transcripts, because they exist in both total cytoplasmic and poly(A)-selected RNA and do not hybridize with a 3' UTR probe (Supplementary Fig. S2c). Cloning and sequencing of these fore-shortened RNAs identified the polyadenylation sites used (Fig. 1e). Thus, an ORF2 sequence placed in the antisense direction inhibits full-length transcript production primarily through premature polyadenylation. Some truncated, polyadenylated transcripts derived from GFPORF2 were also identified but these specific truncations have only a minor function in decreasing the amounts of full-length transcript. Polyadenylated transcripts are expected to be stable. If these shorter transcripts are primarily responsible for decreased full-length RNA amounts they should be detected at much higher concentrations, and they are in GFPORF2AS. After quantifying total hybridization signal and comparison with the control, only 15% of the 'missing' GFPORF2 RNA is accounted for by lower-molecular-mass species, whereas 87% is accounted for in GFPORF2AS (Supplementary Fig. S3). We also experimentally verified that inserting a poly(A) signal upstream of the L1 ORF2 sequence (immediately after GFP) leads to truncated species that can account for the lack of full-length transcript (Supplementary Fig. S2d). Thus, only some of the decrease in sense-strand GFPORF2 RNA is accounted for by premature polyadenylation, and most of the decrease results from another mechanism.

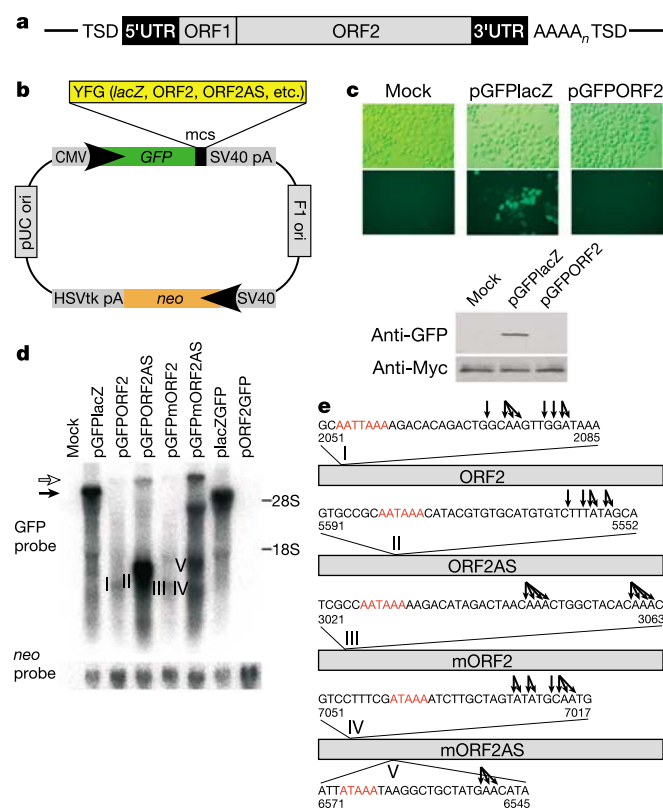


Figure 1 ORF2 sequence decreases expression. **a**, L1 structure (see the text). **b**, Plasmid structures. Fusions are in frame unless noted otherwise. *neo* is used for normalization of transfection. **c**, Immunoblotting of HeLa transfections. Anti-Myc shows equal loading. **d**, Total RNA analysis of HeLa transfections. mORF2, mouse ORF2; AS, antisense. Open and black arrows show the expected positions of GFPORF2 and GFP_{lacZ}, respectively. **e**, Identification of lower-molecular-mass bands. Roman numerals refer to bands in Fig. 1d. The presumptive poly(A) signals are highlighted in red, and the polyadenylation sites are indicated by arrows. Numbers refer to L1.2 sequence (GenBank accession no. M80343) or L1_{spa} sequence (GenBank accession no. AF016099).

Poor expression is not due to ORF2 protein feedback

Poor expression of ORF2 could result from some effect of ORF2 protein. pGFPORF2mut² contains two missense mutations (D205G and D702Y) that destroy the known catalytic activities of ORF2 (refs 3, 4) and alleviate its toxicity in yeast²³ and baculovirus¹². pGFPORF2 was altered separately by adding a stop codon and four extra base pairs between GFP and ORF2; in this construct, GFPstop-ORF2, the GFPORF2 fusion protein is truncated and ORF2 is shifted out of frame with respect to GFP, and produces only GFP. Finally, we introduced two +1 frameshifts early in ORF2 by adding adenosines before nucleotide positions 2,134 and 2,629 (numbers relative to the L1.2 GenBank sequence, accession no. M80343) to eliminate the possibility of the production of ORF2 protein by the initiation of internal translation. Although these frameshifts are early in ORF2, they occur after a residue (N14) required for transposition³, and thus must be translated in a full-length active L1 element, whether or not ORF2 is initiated internally. pGFPORF2, pGFPORF2mut², pGFPstopORF2 and pGFPORF2fs each show similar reductions in RNA (Fig. 2), confirming that the RNA deficit results from the ORF2 nucleotide sequence, not ORF2 protein.

Inhibitory effect does not map to a discrete sequence

We attempted to map a region of the ORF2 sequence responsible for this decrease in RNA. Starting with pGFPstopORF2, progressive amino-terminal and carboxy-terminal deletions, as well as several internal deletions, were tested (Fig. 3a). No single discrete nucleotide sequence block determined the RNA deficit. Rather, longer ORF2 sequences led to lower concentrations of RNA, which correlated with protein abundance (Supplementary Fig. S3). This is consistent with a repetitive sequence or sequences scattered throughout ORF2 that collectively inhibit ORF2 expression.

An interesting feature of L1 coding sequences is a strong adenosine-rich bias in the sense strand (Fig. 3b). This A-rich bias is absent from the 5' UTR. We tested the hypothesis that ORF1 and ORF2 confer a similar length-dependent inhibition of expression. Because ORF1 is only 1,017 base pairs long, placing ORF1 downstream of

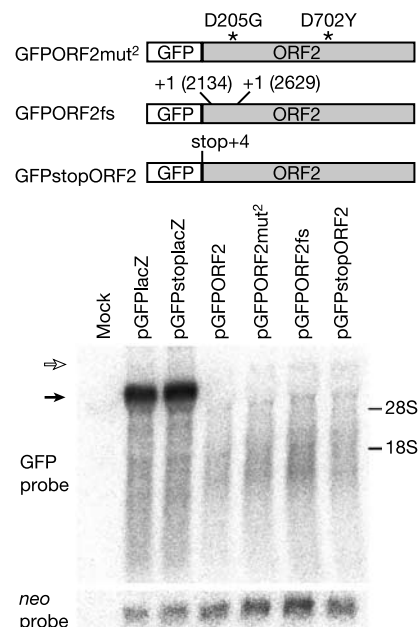


Figure 2 Decreased RNA amounts are not due to ORF2 protein. Top, structures of GFPORF2mut², GFPORF2fs and GFPstopORF2. Bottom, total RNA analysis of HeLa transfections. Open and black arrows show the expected positions of GFPORF2 and GFP_{lacZ}, respectively.

GFP is not expected to interfere with its expression on the basis of the ORF2 deletion mapping experiments. However, if the A-rich bias governs poor expression, expanding the ORF1 sequence to ORF2 length (about 4 kilobases) should markedly decrease RNA amounts. A control construct (pGFPORF1) with a single copy of ORF1 is expressed efficiently, but a lengthened ORF1-containing construct (pGFPstop4ORF1) containing four tandem repeats of ORF1, like the ORF2 construct, led to low concentrations of GFP RNA. As with ORF2, the effect was sense-strand-specific (Fig. 3c, lanes 4–7). In contrast, tandemized L1 5' UTR (5 × 800 base pairs long), lacking the A-rich bias, had no effect (Fig. 3c, lanes 8 and 9). The effect of unusual base composition on transcription might be analogous to a similar phenomenon in yeast, in which mutants in the THO complex decrease the transcription of long GC-rich DNA stretches²⁴. Elsewhere we show that alteration of the L1 base content abolishes the transcription defect and results in high-frequency retrotransposition²⁵, further supporting the hypothesis that the unusual strand bias inhibits expression.

ORF2 transcripts are inefficiently elongated

Lower steady-state RNA concentrations could result from either increased RNA degradation or decreased RNA production. We measured the RNA half-life of the GFPlacZ and GFPORF2 transcripts after treatment with the transcriptional inhibitor actinomycin D. GFPlacZ and GFPORF2 transcript amounts decreased on a similar timescale relative to the *neo* transcript (Fig. 4a, left). The half-life of GFPORF2 relative to GFPlacZ is decreased at most twofold (Fig. 4a, right); quantification by real-time polymerase

chain reaction with reverse transcription (RT-PCR) was consistent with this result (Supplementary Fig. S4). This minor change in RNA half-life cannot account for the 70-fold decrease in steady-state GFPORF2 RNA amounts (Fig. 1d, lanes 2 and 3), indicating that most of the GFPORF2 transcript decrease might result from some mechanism other than transcript instability.

A nuclear run-on assay was performed to evaluate RNA polymerase density along GFPlacZ and GFPmORF2 transcripts produced in human nuclei. Mouse ORF2 was used for this hybridization-based assay to avoid the high background expected from the 500,000 endogenous human L1 elements. These experiments show that ORF2 does not inhibit transcription initiation, because polymerase density in the early region of GFPmORF2 is similar to the same region in GFPlacZ (Fig. 4b, compare GFP1 probes). However, whereas RNA polymerase proceeds through the GFPlacZ transcript processively, functionally engaged RNA polymerase density decreases gradually as transcription is assayed along the ORF2 sequence (Fig. 4b). Thus, the ORF2 sequence seems to be a poor substrate for transcriptional elongation. This could be due to slow elongation rate, stalling of the RNA polymerase complex, or premature dissociation.

Evidence for L1 involvement in transcriptome evolution

Because L1 is a mobile element with relatively non-specific target site selection^{3,12,14,21,22,26}, the observed transcriptional properties of L1 sequences take on potentially great significance. When L1 elements insert into introns, the new L1 sequence inevitably becomes part of the target gene and its transcript. On the basis of our data, we predict that L1 insertions in either orientation could attenuate the expression of target genes by premature truncation of RNA (for antisense insertions) or a transcriptional elongation defect (for sense insertions), each of which decreases the production of full-length pre-messenger RNA. Because the L1 sequence in the sense orientation with respect to transcription of the target gene decreases the production of full-length products to a greater extent than the L1 antisense sequence (Fig. 1d), we expect that this effect would require somewhat more inserted antisense sequence within the target gene. However, decreased target gene expression is expected in both L1 orientations. To investigate these hypotheses, we used expression profiling data to select the 5% most highly and most poorly expressed genes in humans. We then examined the genomic loci of these two sets of genes and used RepeatMasker to search the predicted pre-mRNA transcripts for L1 sequence.

The average total amount of L1 sequence present per gene was markedly different for highly expressed and poorly expressed genes. Highly expressed genes had small amounts of L1 sequence (an average of 918 nucleotides of sense L1 per gene and 1,760 nucleotides of antisense L1 per gene), whereas poorly expressed genes had large amounts of L1 sequence (an average of 4,760 nucleotides of sense L1 per gene and 8,860 nucleotides of antisense L1 per gene; see Fig. 5a). In comparison with randomly selected populations of genes, these values represent the two possible extremes on the spectrum (Fig. 5b). When the total amount of L1 sequence present was normalized for total intron amount, L1 was still underrepresented in highly expressed genes (Fig. 5c), showing that this effect cannot be entirely explained by the greater total intron content of the poorly expressed genes. These data suggest that L1 sequence content is one of many contributing factors that determine the expression of a particular gene. Further analysis shows that this relationship holds even within specific isochores (L1-poor/highly expressed or L1-rich/poorly expressed) of the human genome (Fig. 5d).

Discussion

L1 ORF sequences in the sense orientation serve as a poor substrate for transcription. The inhibition of transcription by the L1 ORFs could well serve as a negative regulatory mechanism evolved by the

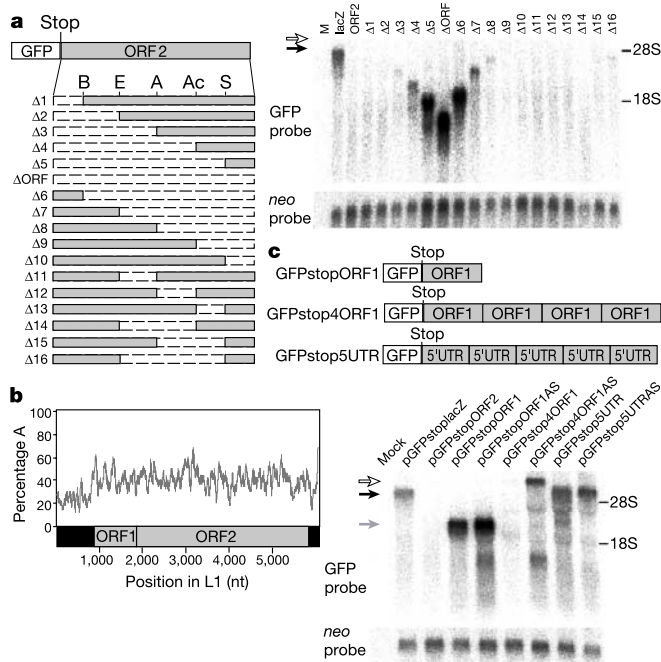


Figure 3 Decrease in L1 expression is dependent on length. **a**, The left panel depicts the structures of deletion constructs. Hollow regions represent deleted sequences. B, *BbvCI*; E, *EcoRI*; A, *AflIII*; Ac, *AcI*; S, *SpeI*. The right panel shows a total RNA analysis of HeLa transfections. Lanes: M, mock; lacZ, pGFPstoplacZ; ORF2, pGFPstopORF2. Open and black arrows show the expected positions of GFPstopORF2 and GFPstoplacZ, respectively. **b**, The adenosine base composition of the sense strand, in 50-nucleotide windows, was plotted for each position in L1.2 with MacVector 6.5.3 (Oxford Molecular). **c**, The top panel shows the structures of GFPstopORF1, GFPstop4ORF1 and GFPstop5UTR. The 4ORF1 repeat is about 4,500 nucleotides long and the 5' UTR repeat is about 4,000 nucleotides long. The bottom panel shows a total RNA analysis of HeLa transfections. Open, black and grey arrows show the expected positions of GFPstop4ORF1, GFPstop5UTR and GFPstopORF1, respectively.

host to prevent excessive retrotransposition. This is consistent with the ability to express L1 ORF2 in non-mammalian organisms that lack the L1 family of retrotransposons and therefore might not have evolved the required regulatory machinery. It is possible that in retrotransposition competent tissues (for example germ cells^{17,18}) a more processive form of the RNA polymerase II complex is recruited to the L1 promoter and that this bypasses the elongation defect. This could be due either to germ-cell- or other cell-specific transcription factors that affect the type of RNA polymerase elongation complex formed at the L1 promoter^{27–29}, or germ-cell-specific elongation factors^{30,31} that enhance active L1 transcription.

The changes in target gene expression and structure that could result from an L1 insertion are potentially of even greater significance. Because RNA polymerase gradually pauses and/or dissociates from the template as it encounters longer stretches of L1 sequence, we expect L1 insertions to attenuate expression of the target gene (Fig. 6a). Bioinformatic data presented here support the hypothesis that L1 insertions attenuate gene expression and have a profound impact on the human transcriptome. Our data are also consistent with examples of known *de novo* full-length L1 insertions into introns that lead to very significantly decreased RNA amounts of target genes in mammalian cells^{32,33}. Further experiments will be needed to determine whether the inhibitory effect is cumulative

over multiple segments of L1 ORF-derived sequence or whether all L1 sequence must be in a single contiguous block to inhibit elongation.

When molecular parasites invade a genome, they evolve to maximize their own survival, often to the detriment of their host^{34,35}. Nevertheless, this unwelcome guest might become useful over evolutionary time, as has occurred with significant consequences in *Drosophila*, in which telomeres are produced by retrotransposition³⁶, and in the evolution of the vertebrate adaptive immune system³⁷. In the present case, L1 elements might have had an integral role in the evolution of humans by repeatedly fine-tuning gene expression. Because L1s are found in high copy number in all mammals, and mouse ORF2 behaves similarly to human ORF2, this proposed model could apply throughout mammalian evolution. The fine-tuning could be a cumulative effect of small L1 insertions into introns that individually might have relatively minor effects on expression. When such expression adjustments are beneficial, the insertions will provide a selective advantage and become fixed in the host genome; when effects are deleterious, the insertion alleles will fail to become fixed in the gene pool. Because large insertions within genes are likely to have an immediate, marked impact on expression of the target, they are more likely to be selected against¹³.

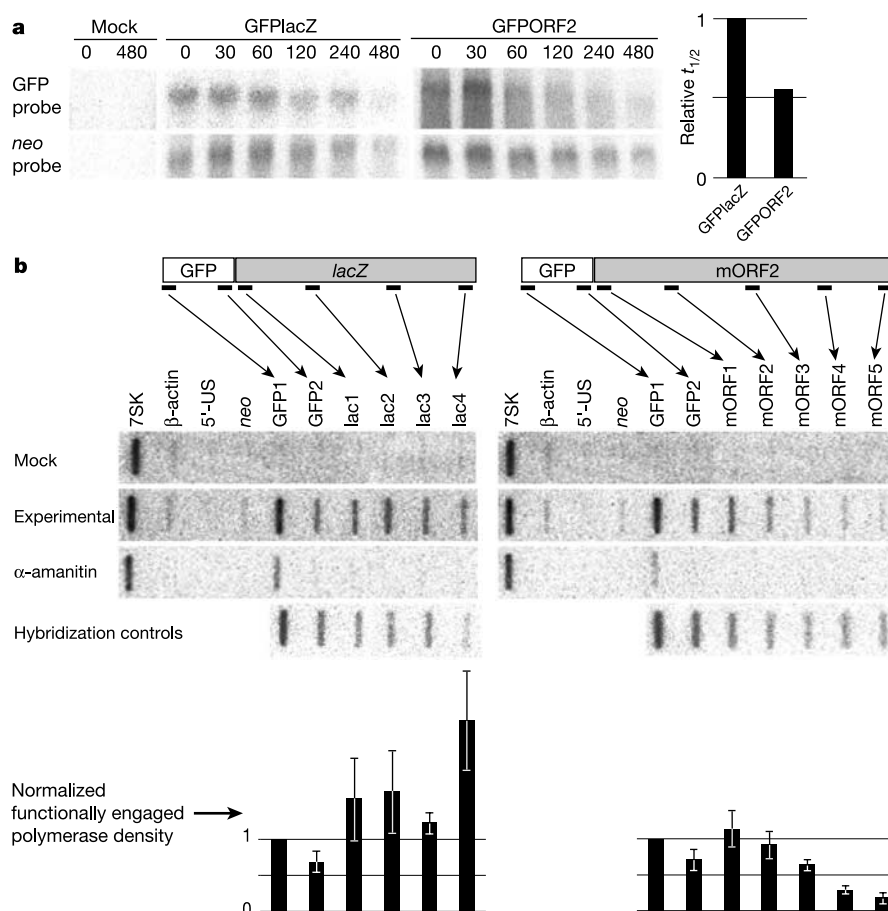


Figure 4 Analysis of ORF2 stability and transcription. **a**, Half-life measurements. In the left panel, transfected HeLa cells were treated with actinomycin D 48 h after transfection. Total RNA was collected 0, 30, 40, 120, 240 and 480 min after treatment and quantified by blotting. In the right panel, the half-lives ($t_{1/2}$) of GFPlacZ or GFPORF2 (relative to the $t_{1/2}$ of the *neo* control) were calculated and compared with $t_{1/2, \text{GFPlacZ}}/t_{1/2, \text{neo}}$ set to 1. **b**, Nuclear run-on analysis (NRO). Nuclei were isolated from HeLa cells 36 h after transfection and used for NRO. Bold lines under GFP, *lacZ* and mORF2 indicate probe

positions. 7SK controls for RNA polymerase III transcription. β -actin is a control for RNA polymerase II transcription. 5'-US is a negative control that hybridizes to a region upstream of the cytomegalovirus (CMV) promoter. *neo* controls for transfection. Hybridization controls are described in Methods. Normalized functionally engaged polymerase density is the signal ($N = 3$) corrected for α -amanitin-resistant transcription and hybridization efficiency, with GFP1 set to 1. Error bars show the standard deviation.

Our findings also indicate that L1 insertions into endogenous genes could potentially generate novel mRNA, and consequently protein isoforms, by producing prematurely truncated, stable transcripts in addition to the original gene product (Fig. 6b). An example of a useful short mRNA/protein isoform generated from an

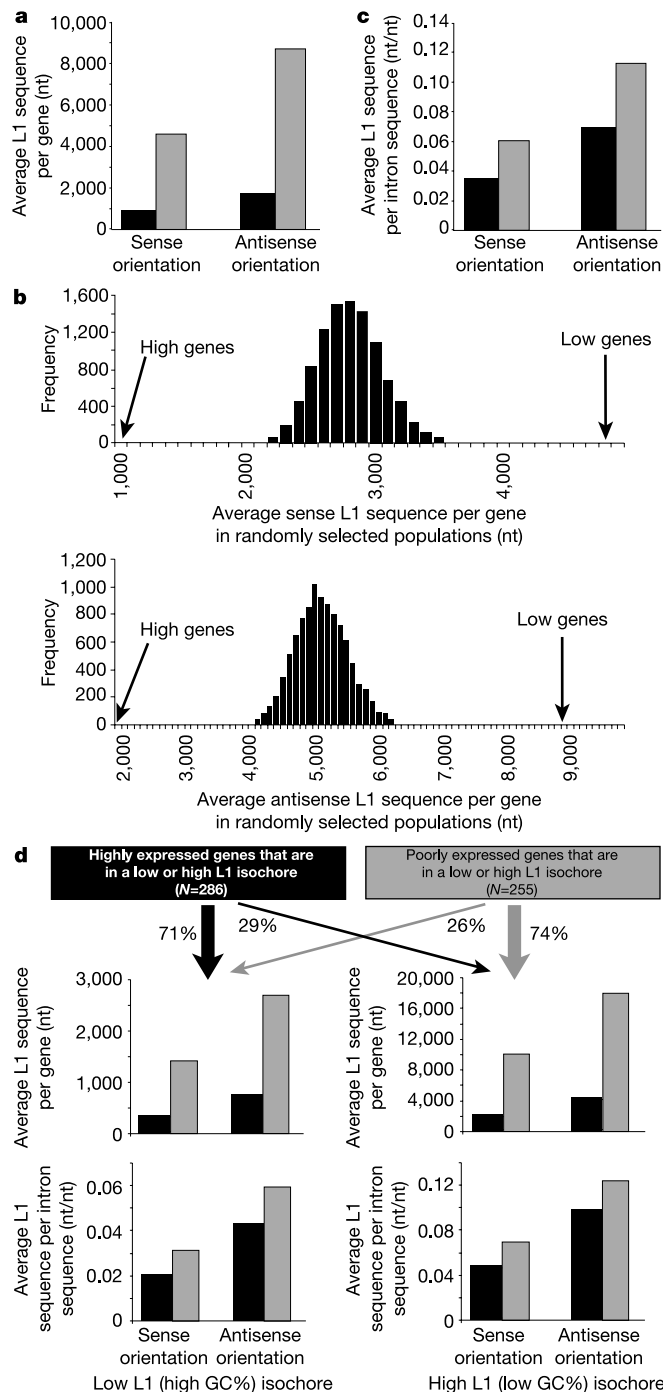


Figure 5 Bioinformatic analysis of L1 content in genes. **a**, Average L1 content of genomic loci of sets of highly (black bars) and poorly (grey bars) expressed genes (see Methods). **b**, Average L1 content in sets of randomly selected populations of genes (see Methods). Positions where the highly and poorly expressed genes would be (data superimposed from **a**) are indicated and are outside the random distribution ($P < 0.01$). **c**, Data from **a**, normalized to total intron content. **d**, Highly and poorly expressed genes were sorted into high GC, low L1 isochore or low GC, high L1 isochore⁵⁰ classes. The percentage of each expression class falling into each isochore is indicated. Subpopulations were analysed as described in **a** and **c**.

L1 insertion has been shown for the human ATRN gene³⁸. In this example, the L1 insertion is short (212 base pairs) and the polyadenylation signal used is the presumed native L1 polyadenylation signal. We predict that truncated mRNA isoforms can be produced from other polyadenylation signals in L1 coding sequences, particularly in the L1 antisense orientation. We have found examples of human transcripts that use precisely these cryptic poly(A) signals (for example, SPC25 gene transcripts, some of which end at poly(A) signal II from Fig. 1d) (S. Wheelan and J.D.B., manuscript in preparation). The cryptic poly(A) signals seem to occur at sites scattered throughout the element sequence. All of these contain the highly conserved AATAAA element or a subtle variant (AATTAAA or ATAAA) at the appropriate distance from the site of cleavage and polyadenylation (Fig. 1e). The downstream GU/U-rich regions of poly(A) sites are highly variable³⁹, and we suspect that the 40% U content of the ORF2 antisense strand might account for the stronger polyadenylation signals observed in the antisense case. The weakness of the sites in the sense strand presumably reflects ongoing selection on L1 to make full-length RNA copies for retrotransposition. Premature polyadenylation of L1 sense sequences has recently been noted by others¹⁹ and is consistent with this aspect of the proposed model.

These models (Fig. 6a, b) predict that L1 insertions could have major effects on both the quality and quantity of genome-wide mRNAs. Longer L1 insertions should have more potent effects and are therefore more likely to be disastrous for target gene expression. The potential danger of long L1 insertions is consistent with the observed profile of L1 insertions in humans, in which short segments corresponding to 3'-terminal fragments of the L1 element dominate the length distribution^{14,20–22}. This is probably due to a combination of non-processive reverse transcription and the sub-

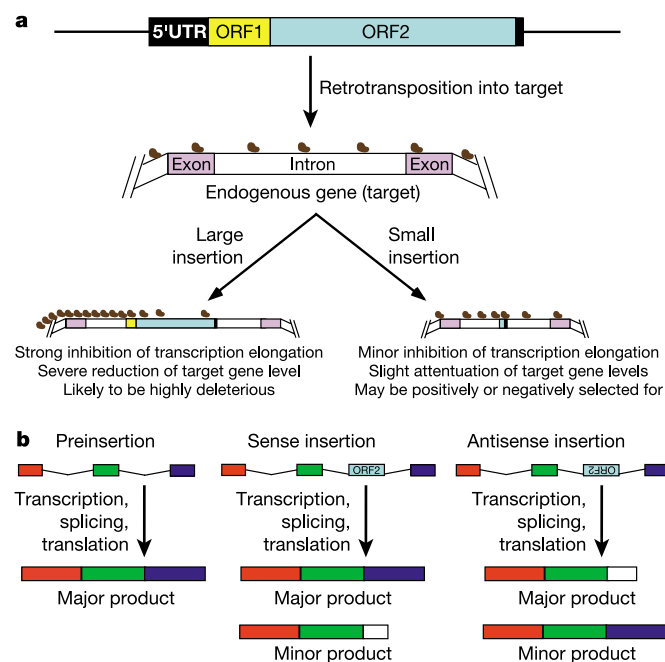


Figure 6 Models for L1-mediated modulation of gene expression/structure. **a**, Effects on transcription. Brown dots represent transcriptional complexes, which could be slowed, paused or dissociated from the templates on encountering significant amounts of L1 sequence. **b**, Effects on mRNA and protein structure. Left, hypothetical gene with three exons. Middle, intronic sense L1 insertions can produce a minor amount of prematurely polyadenylated mRNA, potentially giving rise to a truncated protein with additional, previously untranslated amino acids at the C terminus (white segment). Right, intronic antisense L1 insertions can produce a major amount of prematurely polyadenylated mRNA.

sequent 'purifying selection' against long insertions¹³. In addition, if L1 functions to fine-tune gene expression, retrotransposition-competent (and by definition full-length) L1 elements could provide a selective advantage to the host and might be present at a higher than expected frequency. This might help to explain a puzzling aspect of the L1 insertion profile—the overrepresentation of full-length L1 elements^{13,14}. Although many such elements are currently non-functional because of mutation, most were probably previously active.

L1 elements and their kin in other eukaryotes have been shown to rearrange genes and chromosomes by a wide variety of mechanisms. In addition to gene inactivation through insertion into exons, insertions in tissue culture cells are often associated with genomic instability^{21,22}. Homologous recombination between dispersed copies of retrotransposon sequences can also lead to chromosome rearrangements⁴⁰. Finally, transcription past the L1 element polyadenylation site during retrotransposition can lead to the 'transduction' of 3' sequences flanking the donor L1 element^{41–43}. A symmetrical process involving 5' sequences can occur when a cellular promoter upstream of an L1 element transcribes into an active L1 element²². Here we propose an addition to the retrotransposon repertoire of genome remodelling activities—the insertion of expression-modulating sequences into host gene introns to reprogram gene expression. As recent genome sequencing studies have shown, the gene complements of mammals can be remarkably similar⁴⁴. Thus, a major component of mammalian speciation might result from subtle transcriptome reprogramming through alterations in exon usage patterns as well as amounts of gene expression. Even within a species, alleles that show *cis*-acting heritable variations in expression are relatively common in normal individuals^{45,46} and might account for phenotypic differences. The proposed model suggests the possibility that L1 elements could have a major function in such variations, both within and between species. Rigorous experimental proof of this model requires knocking L1 sequences into the introns of specific mammalian genes, generating isogenic filled and empty alleles, and comparing the quantity and quality of the resultant gene products. □

Methods

Oligonucleotide sequences, plasmid construction

All oligonucleotide sequences described in this manuscript are shown in Supplementary Table S1. General structures of test expression reporters are shown (Supplementary Fig. S1); details of plasmid construction are available from the authors on request.

Cell culture and transfection

Cell culture and transfections were performed as described²⁵. Proportions were scaled up linearly for 150-mm dishes. For downstream northern, immunoblot or nuclear run-on analysis, cells were harvested 36–48 h after transfection.

Northern blot analysis

Northern blots were performed essentially as described²⁵. Prehybridizations and hybridizations were performed in 50% formamide, 5× SSC, 5× Denhardt's solution, 1% SDS, and 100 µg ml^{−1} boiled herring-sperm DNA at 42 °C. The following [γ -³²P]ATP end-labelled oligonucleotides were used as probes: GFP probe, JB4057; *neo* probe, JB4059; 3' UTR probe, JB5574.

Image Gauge v. 3.0 (Fuji Photo Film Co.) was used to quantify the signal ($N = 3$). Each band was corrected by subtracting a lane-specific background. Corrected test transcripts were normalized to corrected *neo* transcripts.

For half-life measurements ($N = 2–4$), actinomycin D was used at 5 µg ml^{−1}. A radiolabelled 500-base-pair fragment of the enhanced GFP gene was used to detect GFP fusion transcripts. End-labelled JB4059 was used to detect *neo* transcripts. GFPORF2 samples were exposed 50-fold longer than GFP Δ to make them visible. The quantified transcripts, normalized to *neo*, were plotted on a semi-logarithmic axis against time, where the slope from a best-fit line represents the decrease in relative transcript over time. The slope from these plots of GFP Δ and GFPORF2 were compared to obtain the relative half-lives shown in Fig. 4a (right).

Western blot analysis

Immunoblots were performed as described elsewhere²⁵. Anti-GFP(FL) antibody (Santa Cruz) was used at 1:500 dilution. Anti-rabbit IgG (Amersham) was used at 1:5,000 dilution. Anti-Myc antibody (Santa Cruz) was used at 1:1,000 dilution. Anti-mouse IgG (Amersham) was used at 1:5,000 dilution.

Fluorescence microscopy

Cells were visualized with a Nikon Eclipse TE300 microscope using a fluorescein isothiocyanate filter (Chroma B-2E/C) for GFP visualization. The same exposure time was used for all fluorescence pictures.

RT-PCR cloning

First-strand synthesis was performed with SUPERScript II RNaseH[−] Reverse Transcriptase (Invitrogen) in accordance with the manufacturer's instructions. For each first-strand synthesis reaction, 40 ng DNase I-treated total RNA was used. The primer 3RACERT was used for first-strand synthesis. A 1-µl portion of the first-strand synthesis reaction was used in a PCR reaction containing 100 mM dNTPs, each primer at 500 nM, and 5 units of AmpliTaq (Perkin Elmer). Amplification primers were 3RACEAMP and JB4360. Amplification products were separated on a 1% agarose gel and dominant bands were excised and cloned using the TOPO TA cloning kit (Invitrogen). DNA sequencing was performed by Agencourt.

Nuclear run-on analysis

Isolation and transcription of nuclei. Nuclei were isolated and run-on reactions were performed essentially as described previously⁴⁷. Reactions were stopped by the addition of TRIzol, and RNA was isolated in accordance with the manufacturer's instructions. RNA was hydrolysed and hybridized as described previously⁴⁷.

Filter preparation/hybridization. Each oligonucleotide (0.75 pmol) was diluted in 100 µl of 0.5 M NaOH and slot-blotted on a GeneScreen Plus nylon membrane. Filters were neutralized with 100 mM Tris-HCl, pH 8, crosslinked by ultraviolet irradiation, then prehybridized and hybridized in 50% formamide, 5× SSC, 5× Denhardt's solution, 1% SDS and 100 µg ml^{−1} boiled herring-sperm DNA.

Control hybridization transcripts. A T7 transcription reaction of each control transcript with a C-terminal ATP-binding aptamer⁴⁸ was performed *in vitro* with radiolabelled CTP. Full-length transcripts were purified on an ATP-agarose column as described previously⁴⁹. These transcripts were hydrolysed and hybridized as described above.

Signal quantification. Each quantified slot-blot signal was corrected for membrane background by subtracting signal from an adjacent region of the membrane. Each slot-blot signal was also corrected for α -amanitin-resistant transcription. To measure elongation, the corrected experimental signal was divided by the hybridization correction factor (background-corrected control hybridization signal); results were then normalized to GFP1 by dividing all signals by the GFP1 signal. To measure initiation, the corrected GFP1 experimental signal was divided by the corresponding *neo* experimental signal.

Bioinformatic analysis of L1s in genes

The gene expression data of normal tissues profiled by Gene Logic Inc. were analysed to identify 983 genes with the highest expression and 866 genes with the poorest expression. The genomic coordinates (transcript start to transcript end) of 16,597 RefSeq genes and the genome RepeatMasker*.out files were downloaded from <http://genome.ucsc.edu/> (June 2002 genome assembly, build 12). The RepeatMasker*.out files were parsed to identify the genomic coordinates of L1 elements. These L1 coordinates were used to calculate the amount of 'LINE1/L1' in each transcript.

The distributions of L1s in random genes were constructed by performing bootstrapping. In brief, a set of genes was randomly generated from the 16,597 RefSeq genes. The L1 content was recorded for each random gene set. This process was repeated for 10,000 iterations to generate a distribution.

Isochores⁵⁰ with low (40% or less) GC content and high (50% or more) GC content were obtained from <http://bioinfo2.ugr.es/isochores/>. The highly and poorly expressed genes belonging to these subsets were analysed for L1 content as described above. Genes spanning isochore boundaries were added to both isochores (this was rare). Some genes did not fall into either isochore subset and were therefore omitted from this analysis.

Received 22 January; accepted 30 March 2004; doi:10.1038/nature02536.

1. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
2. Swergold, G. D. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.* **10**, 6718–6729 (1990).
3. Feng, Q., Moran, J. V., Kazanian, H. H. Jr & Boeke, J. D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916 (1996).
4. Moran, J. V. *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917–927 (1996).
5. Esnault, C., Maestre, J. & Heidmann, T. Human LINE retrotransposons generate processed pseudogenes. *Nature Genet.* **24**, 363–367 (2000).
6. Wei, W. *et al.* Human L1 retrotransposition: *cis* preference versus *trans* complementation. *Mol. Cell. Biol.* **21**, 1429–1439 (2001).
7. Kolosha, V. O. & Martin, S. L. *In vitro* properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. *Proc. Natl Acad. Sci. USA* **94**, 10155–10160 (1997).
8. Martin, S. L. & Bushman, F. D. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Mol. Cell. Biol.* **21**, 467–475 (2001).
9. Kolosha, V. O. & Martin, S. L. High-affinity, non-sequence-specific RNA binding by the open reading frame 1 (ORF1) protein from long interspersed nuclear element 1 (LINE-1). *J. Biol. Chem.* **278**, 8112–8117 (2003).
10. Dewannieux, M., Esnault, C. & Heidmann, T. LINE-mediated retrotransposition of marked Alu sequences. *Nature Genet.* **35**, 41–48 (2003).
11. Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**, 595–605 (1993).

12. Cost, G. J., Feng, Q., Jacquier, A. & Boeke, J. D. Human L1 element target-primed reverse transcription *in vitro*. *EMBO J.* **21**, 5899–5910 (2002).
13. Boissinot, S., Entezam, A. & Furano, A. V. Selection against deleterious LINE-1-containing loci in the human lineage. *Mol. Biol. Evol.* **18**, 926–935 (2001).
14. Szak, S. T. *et al.* Molecular archeology of L1 insertions in the human genome. *Genome Biol.* **3**, research00521–research005218 (2002).
15. Mathias, S. L., Scott, A. F., Kazazian, H. H. Jr, Boeke, J. D. & Gabriel, A. Reverse transcriptase encoded by a human transposable element. *Science* **254**, 1808–1810 (1991).
16. Clements, A. P. & Singer, M. F. The human LINE-1 reverse transcriptase: effect of deletions outside the common reverse transcriptase domain. *Nucleic Acids Res.* **26**, 528–535 (1998).
17. Branciforte, D. & Martin, S. L. Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Mol. Cell. Biol.* **14**, 2584–2592 (1994).
18. Trelogan, S. A. & Martin, S. L. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. *Proc. Natl Acad. Sci. USA* **92**, 1520–1524 (1995).
19. Perepelitsa-Belancio, V. & Deininger, P. RNA truncation by premature polyadenylation attenuates human mobile element activity. *Nature Genet.* **35**, 363–366 (2003).
20. Boissinot, S., Chevret, P. & Furano, A. V. L1 (LINE-1) retrotransposon evolution and amplification in recent human history. *Mol. Biol. Evol.* **17**, 915–928 (2000).
21. Gilbert, N., Lutz-Prigge, S. & Moran, J. V. Genomic deletions created upon LINE-1 retrotransposition. *Cell* **110**, 315–325 (2002).
22. Symer, D. E. *et al.* Human L1 retrotransposition is associated with genetic instability *in vivo*. *Cell* **110**, 327–338 (2002).
23. Feng, Q. *Mechanism of Human L1 Element Retrotransposition*. Thesis, Johns Hopkins Univ. (1996).
24. Chávez, S. *et al.* Hpr1 is preferentially required for transcription of either long or G + C-rich DNA sequences in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **21**, 7054–7064 (2001).
25. Han, J. S. & Boeke, J. D. A highly active synthetic mammalian retrotransposon. *Nature* **429**, 314–318 (2004).
26. Cost, G. J. & Boeke, J. D. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochemistry* **37**, 18081–18093 (1998).
27. Minakami, R. *et al.* Identification of a cis-element essential for the human L1 transcription and a nuclear factor(s) binding to the element. *Nucleic Acids Res.* **20**, 3139–3145 (1992).
28. Tchenio, T., Casella, J. F. & Heidmann, T. Members of the SRY family regulate the human LINE-1 retrotransposons. *Nucleic Acids Res.* **28**, 411–415 (2000).
29. Yang, N., Zhang, L., Zhang, Y. & Kazazian, H. H. Jr An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucleic Acids Res.* **31**, 4929–4940 (2003).
30. Xu, Q., Nakanishi, T., Sekimizu, K. & Natori, S. Cloning and identification of testis-specific transcription elongation factor S-II. *J. Biol. Chem.* **269**, 3100–3103 (1994).
31. Miller, T., Williams, K., Johnstone, R. W. & Shilatfard, A. Identification, cloning, expression, and biochemical characterization of the testis-specific RNA polymerase II elongation factor ELL3. *J. Biol. Chem.* **275**, 32052–32056 (2000).
32. Schwahn, U. *et al.* Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nature Genet.* **19**, 327–332 (1998).
33. Yajima, I. *et al.* An L1 element intronic insertion in the black-eyed white (Mitf^{mi-bw}) gene: the loss of a single Mitf isoform responsible for the pigmentary defect and inner ear deafness. *Hum. Mol. Genet.* **8**, 1431–1441 (1999).
34. Hickey, D. A. Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics* **101**, 519–531 (1982).
35. Bestor, T. H. Sex brings transposons and genomes into conflict. *Genetica* **107**, 289–295 (1999).
36. Levis, R. W., Ganesan, R., Houtchens, K., Tolar, L. A. & Sheen, F. M. Transposons in place of telomeric repeats at a *Drosophila* telomere. *Cell* **75**, 1083–1093 (1993).
37. Agrawal, A., Eastman, Q. M. & Schatz, D. G. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* **394**, 744–751 (1998).
38. Tang, W. *et al.* Secreted and membrane attractin result from alternative splicing of the human ATRN gene. *Proc. Natl Acad. Sci. USA* **97**, 6025–6030 (2000).
39. Zarudnaya, M. I., Kolomiets, I. M., Potyahaylo, A. L. & Hovorun, D. M. Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures. *Nucleic Acids Res.* **31**, 1375–1386 (2003).
40. Burwinkel, B. & Kilimann, M. W. Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *J. Mol. Biol.* **277**, 513–517 (1998).
41. Moran, J. V., DeBerardinis, R. J. & Kazazian, H. H. Jr Exon shuffling by L1 retrotransposition. *Science* **283**, 1530–1534 (1999).
42. Goodier, J. L., Ostertag, E. M. & Kazazian, H. H. Jr Transduction of 3′-flanking sequences is common in L1 retrotransposition. *Hum. Mol. Genet.* **9**, 653–657 (2000).
43. Pickeral, O. K., Makalowski, W., Boguski, M. S. & Boeke, J. D. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. *Genome Res.* **10**, 411–415 (2000).
44. Okazaki, Y. *et al.* Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420**, 563–573 (2002).
45. Yan, H., Yuan, W., Velculescu, V. E., Vogelstein, B. & Kinzler, K. W. Allelic variation in human gene expression. *Science* **297**, 1143 (2002).
46. Lo, H. S. *et al.* Allelic variation in gene expression is common in the human genome. *Genome Res.* **13**, 1855–1862 (2003).
47. Cuello, P., Boyd, D. C., Dye, M. J., Proudfoot, N. J. & Murphy, S. Transcription of the human U2 snRNA genes continues beyond the 3′ box *in vivo*. *EMBO J.* **18**, 2867–2877 (1999).
48. Dieckmann, T., Butcher, S. E., Sassanfar, M., Szostak, J. W. & Feigon, J. Mutant ATP-binding RNA aptamers reveal the structural basis for ligand binding. *J. Mol. Biol.* **273**, 467–478 (1997).
49. Sassanfar, M. & Szostak, J. W. An RNA motif that binds ATP. *Nature* **364**, 550–553 (1993).
50. Oliver, J. L. *et al.* Isochore chromosome maps of the human genome. *Gene* **300**, 117–127 (2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank Y. Aizawa, J. Corden, J. Moran, J. Nathans, S.-L. Ooi and D. Valle for helpful discussions and critical reading of the manuscript, S. Wheelan for unpublished bioinformatics data, S. Murphy for providing a detailed nuclear run-on protocol, J. Moran for pY104, H. Kazazian for pTN201, and R. Bandaru for help with statistical analysis. This work was supported by the NIH (J.D.B.) and the Medical Scientist Training Program (J.S.H.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.D.B. (jboeke@jhmi.edu).

An age–colour relationship for main-belt S-complex asteroids

Robert Jedicke¹, David Nesvorný², Robert Whiteley³, Željko Ivezić⁴
& Mario Jurić⁴

¹Institute for Astronomy, University of Hawaii, Honolulu, Hawaii 96822, USA

²Southwest Research Institute, 1050 Walnut St, Suite 400, Boulder, Colorado 80302, USA

³USAF Space and Missile Systems Center, 2420 Vela Way, Suite 146, Los Angeles Air Force Base, El Segundo, California 90245, USA

⁴Princeton University, Department of Astrophysical Sciences, Princeton, New Jersey 08544, USA

Asteroid collisions in the main belt eject fragments that may eventually land on Earth as meteorites^{1–3}. It has therefore been a long-standing puzzle in planetary science that laboratory spectra of the most populous class of meteorite (ordinary chondrites, OC) do not match the remotely observed surface spectra of their presumed (S-complex) asteroidal parent bodies. One of the proposed solutions to this perplexing observation is that ‘space weathering’ modifies the exposed planetary surfaces over time through a variety of processes (such as solar and cosmic ray bombardment, micro-meteorite bombardment, and so on). Space weathering has been observed on lunar samples⁴, in Earth-based laboratory experiments^{5,6}, and there is good evidence from spacecraft data that the process is active on asteroid surfaces^{7,8}. Here, we present a measurement of the rate of space weathering on S-complex main-belt asteroids using a relationship between the ages of asteroid families and their colours⁹. Extrapolating this age–colour relationship to very young ages yields a good match to the colour of freshly cut OC meteorite samples, lending strong support to a genetic relationship between them and the S-complex asteroids.

Neither laboratory nor spacecraft measurements have measured the rate of space weathering on asteroids. In the laboratory it is difficult to calibrate the techniques used in simulating the process (for example, rapid laser heating of crushed OC meteorite samples^{5–6}) with the presumed actual mechanism (micro-meteoroid-induced heating of regolith grains). Estimates of characteristic time scales^{6,10} for asteroidal space weathering based on laboratory results vary from 5×10^4 yr to 10^8 yr. From the spacecraft’s vantage point it is possible to distinguish young from old terrains on the basis of crater counting or other geological morphology (for example, rock slides) but it is difficult to date the surface’s absolute ages, because cratering rates on asteroid surfaces vary with time, orbital elements, size and so on. Space weathering on Eros seems to reach maturity on timescales of several tens of millions of years (C. R. Chapman, unpublished work).

We have developed a technique for measuring the rate of space weathering on main-belt asteroids (Fig. 1). Our method uses asteroid families that are genetically related groups of objects that were created in the collisional disruption of a larger asteroid. The surfaces of asteroids in a family must all have the same age as they are all ‘re-set’ by the catastrophic disruption that generated the group. When the collision takes place, many large fragments are ejected and they re-accumulate a layer of material on their surfaces that was formerly in the interior of the parent asteroid¹¹. The identification of asteroid families and determination of their ages is described in Supplementary Information.

We obtained asteroid colours from the Sloan Digital Sky Survey Moving Object Catalog⁹ (SDSS MOC) that contains 5-band (u,g,r,i,z) photometric data accurate to 0.02 magnitudes for over 135,000 entries. More than 28,000 were linked to known asteroids with proper elements from the AstDys database¹². Because our

identification of asteroid families was performed on the same set of proper elements, we were able to cross-link 8,416 asteroids with a known family identification to the MOC. Twelve of the families are classified unambiguously as S-complex by other researchers and/or are well determined by the colours from the MOC itself as described in Supplementary Information.

Although the classification of our families as members of the S-complex is well determined, there are many subtypes within this broad taxonomic complex. In particular, a classification scheme based on mineralogical absorption features¹³ in the infrared suggests that only one subclass (SIV) is mineralogically consistent with an OC meteorite connection. We will argue below that our results suggest that there is a space-weathering connection between S-complex asteroids and the OC meteorites, and it would therefore seem that this analysis should be restricted to only those families of the SIV sub-type. There are two problems with this option: (1) the mineralogical classification scheme has been applied to only 39 asteroids, and only two of those objects are among the thousands of asteroids in our family members, and (2) that classification scheme ignored the possibility of space weathering, assigning all the variability within the S-complex to mineralogical diversification. Similarly, the Eos family contains many K-type (a subclassification within the S-complex) asteroids that are thought to be related to the CO3 or CV3 chondrites rather than the OC meteorites¹⁴.

This study shows that there is an age-dependent component to asteroid colours along with mineralogical differences. As we have no *a priori* means of disentangling these effects, we make no distinction between families of different types within the S-complex. Furthermore, in the wavelength range of the SDSS filters, there are few features available to distinguish quantitatively between S-complex asteroids except for the slope that is related to the first principal component (PC_1) colour shown in Fig. 1.

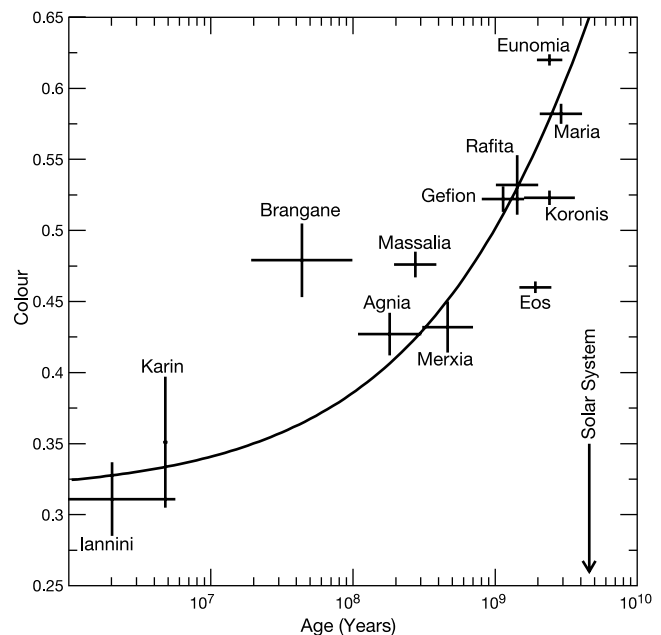


Figure 1 The relationship between the mean colour of S-complex asteroid families and their ages. The PC_1 colour captures $\sim 73\%$ of the variance in colours between asteroids with good spectrophotometric measurements ($\Delta PC_1 < 0.1$) in the SDSS MOC. It is a good proxy for the mean spectral slope between 350 nm and 900 nm. Larger values of PC_1 correspond to higher slopes (redder objects) in that wavelength range. The PC_1 for a family is the standard weighted mean of all measured values for the members of the family. The displayed error bars on PC_1 are the standard error of the weighted mean colour. The solid line represents a fit to the data points accounting for errors in both the ordinate and abscissa.

We used u–g, g–r, g–i and g–z colour bands for 10,776 objects culled from the entire MOC with photometric errors in each band of less than 0.10 magnitudes as input to a principal components (PC) analysis¹⁵ that creates mutually orthogonal linear combinations (eigenvectors) of the colours. In doing so it collapses the input data into successive components that incorporate the maximum remaining variance of the colours. It thus provides a measure of the true dimensionality of the input system so that we can concentrate our efforts on the most important colour terms. We found that ~73% of the variation in the colours is incorporated into $PC_1 = 0.3560m_{ug} + 0.5418m_{gr} + 0.5683m_{gi} + 0.5066m_{gz}$, where m_{xy} is the x–y colour for the asteroid after correction for solar colours [$m_{xy} = (m_x - m_y)_{SDSS} - (m_x - m_y)_{Sun}$]. We have found that PC_1 is well correlated with the slope of the reflectance spectrum in the wavelength range of 350 nm to 900 nm, and is thus a good measure of the reddening effect caused by space weathering. PC_1 is also well correlated with the SDSS colour components¹⁶.

Figure 1 shows that PC_1 depends on a family's age. Assuming that the change in colour is related to the amount of asteroid surface processed by a constant flux of a space-weathering agent, we expect the change of colour as a function of time to be given by $PC_1(t) = PC_1(0) + \Delta PC_1 \{1 - \exp[-(t/\tau)]\}$. We have generalized the form to $PC_1(t) = PC_1(0) + \Delta PC_1 \{1 - \exp[-(t/\tau)^\alpha]\}$ for the fit shown on the figure with $PC_1(0) = 0.32 \pm 0.01$, $\Delta PC_1 = 1.0 \pm 0.1$, $\tau = (25 \pm 5) \times 10^9$ years and $\alpha = 0.50 \pm 0.05$.

The characteristic timescale for space-weathering is $\tau = 10^{10.4 \pm 0.2}$ years, which is two orders of magnitude larger than estimates ($\sim 10^8$ years) based on laboratory measurements of laser-induced simulated weathering of ground meteoritic samples⁶ intended to simulate the effects of micro-meteoroid bombardment. Interaction of the solar wind with asteroid surfaces is predicted¹⁰ to

alter them much faster, perhaps within a characteristic timescale of only $\sim 10^{4.7}$ years.

The caveat is that with existing data our technique is insensitive to measurements of subcomponents of the space-weathering process (for example, there might exist a fast and slow mechanism) if it occurs on timescales of $< \sim 10^6$ years because it would require a sample of very young families (for example, 10^2 – 10^5 years) with well-measured colours. Other known effects of space weathering include spectrum-wide darkening and a reduction in the band depths of absorption features. We were not able to identify any significant change as a function of time in the SIMPS¹⁷ albedo or 2MASS J/H and K/H infrared-band ratios¹⁸ for the S-complex families. This might be due to a lack of sensitivity of our technique (for example, space weathering occurring too fast) and could indicate that different processes act at varying timescales to create the overall space-weathering phenomenon. It is also possible that gardening of the surface regolith through regular bombardment by small asteroids could significantly lengthen the effective space-weathering timescale¹⁰ on the basis of laboratory experiments.

Meteorite spectra in the laboratory are obtained from freshly exposed (ground or crushed) surfaces and therefore correspond to the surface ages of newly formed asteroid families. We predict that the mean PC_1 for new families should be 0.32 ± 0.01 , whereas families formed at the beginning of the Solar System $\sim 4.650 \times 10^9$ years ago should have $PC_1 = 0.67 \pm 0.07$, as shown on Fig. 2. The age–colour correlation presented here, along with a contribution from intrinsic mineralogical differences between the parent bodies, can explain the observation of colour diversity between families¹⁶ within the S-complex.

The mean PC_1 for the OC meteorites in Fig. 2 is 0.39 ± 0.10 (the mean is 0.36 ± 0.06 without the LL3 subclass—with $PC_1 \approx 0.59$ its colours are already similar to the S-complex's). Our prediction for the unweathered colour of S-complex asteroids agrees well with the observed mean PC_1 for the OC meteorites. Our prediction for S-complex asteroid surfaces with an age equal to that of the Solar System is slightly higher than the mean colour of all S-complex asteroids, as would be expected if the majority have experienced some space-weathering. The mode colour of the S-complex asteroids (0.557 ± 0.004) corresponds to an asteroid that has aged for $\sim 1.9 \times 10^9$ years.

Our result, along with a growing body of indirect evidence, strongly supports earlier speculation⁷ that most S-complex asteroids could be OC meteorite parent bodies. Asteroid spectra with typical OC meteorite features are known as Q-types, as exemplified by (1862) Apollo. Even though there is little to no remote-sensing evidence⁷ of Q-type asteroids in the main belt, all four S-complex asteroids for which *in situ* measurements exist (Gaspra, Ida, Dactyl, Eros) have surfaces that are more OC-like in recently excavated terrain leading to the conclusion (C. R. Chapman, unpublished work) that all four are probably OC bodies.

Assuming that S-complex asteroids are the OC parent bodies, our determination of the space-weathering rate correctly predicts the colour of fresh ordinary chondritic material. We predict that as spectroscopists are able to image smaller main-belt asteroids they will increasingly find them to be more Q- or OC-like. Indeed the SMASS2 survey¹⁹ finds some main-belt spectra verging towards Q-type. However, even larger main-belt S-complex asteroids such as (6) Hebe, often cited as a possible OC parent body, can produce OC meteorites. Observations of rotationally resolved spectroscopic heterogeneity²⁰ on that asteroid may be the result of varying degrees of space weathering on different hemispheres.

Given the rate at which main-belt asteroid families are produced, and the distribution of dynamical times for evolution onto Earth-crossing orbits, our results can predict the ratio of Q- to S-type asteroids in the main belt and amongst the near-Earth object (NEO) population. Main-belt S-type asteroids will produce S-type progeny whose freshly exposed surfaces should exhibit Q- or OC-like spectra

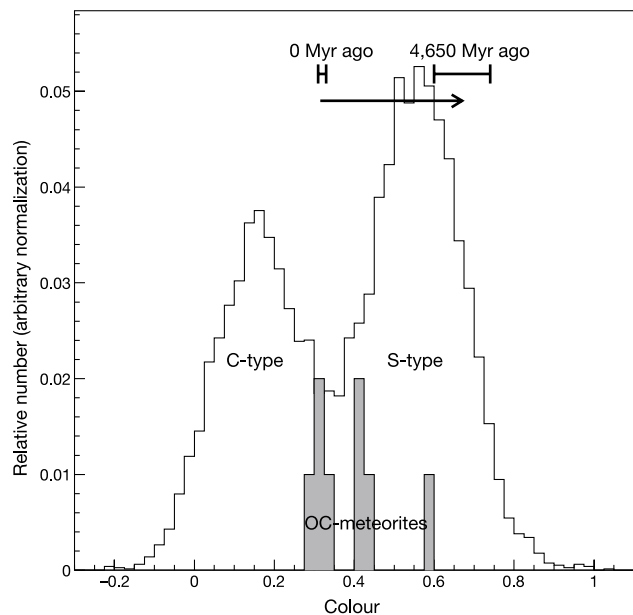


Figure 2 The distribution of colours for S- and C-complex asteroids detected by the SDSS and the average colours of eight types of ordinary chondrite meteorites. Normalized PC_1 distribution for SDSS MOC entries with $\Delta PC_1 < 0.1$ (solid line) and OC meteorites (filled histogram). The peak on the right is the S-complex of asteroids and the peak on the left the C-complex. A gaussian fit to the S-complex PC_1 (> 0.45) distribution yields the most likely S-complex $PC_1 = 0.557 \pm 0.004$. We used the average OC spectra for (from left to right in the histogram) H5, L6 (type i), LL6, H6, L4, L6 (type ii), L5 and LL3 meteorites of ref. 27 convolved with the solar spectrum and the SDSS system transmission then transformed to PC_1 . PC_1 ranges indicated by error bars at the top are predictions for new and old ($4,650 \times 10^6$ years) S-complex material. The arrow points from the central prediction for each in the direction of increased ageing.

after cratering events and catastrophic disruptions. The newly created object's orbits will gradually evolve and perhaps be transported into one of the strong resonances that can pump the orbit's eccentricity to Earth-crossing values. Smaller objects will migrate faster under the influence of the Yarkovsky effect²¹, and once in a resonance the dynamical lifetime of objects of any size is of the order of only a few million years². However, cosmic ray exposure ages tell us that most OC meteorites are tens of millions of years old²². This indicates that even the smallest meteoroids spend a considerable amount of time migrating from their point of origin into the resonances, and the larger objects that eventually become NEOs probably require even more time. Thus, we expect that NEOs of S-complex provenance will display a size-dependent range of spectra ranging from Q (or OC) to S. This prediction is supported by numerous reports in the past decade indicating Q-like spectra and a size-dependent trend to OC-like spectra with decreasing size in the NEO population^{23–26}. However, space weathering in the main belt occurs faster than the lifetimes of NEOs, as suggested by meteoritic cosmic ray exposure ages²². Thus, observational evidence for a large number of OC-like spectra amongst the NEOs^{23–26} implies that regular re-surfacing of these objects keep them looking younger longer. □

Received 21 October 2003; accepted 7 April 2004; doi:10.1038/nature02578.

1. Vokrouhlický, D. & Farinella, P. Efficient delivery of meteorites to the Earth from a wide range of asteroid parent bodies. *Nature* **407**, 606–608 (2000).
2. Gladman, B. J. *et al.* Dynamical lifetimes of objects injected into asteroid belt resonances. *Science* **277**, 197–201 (1997).
3. Wisdom, J. Meteorites may follow a chaotic route to earth. *Nature* **315**, 731–733 (1985).
4. Adams, J. B. & McCord, T. B. Alteration of lunar optical properties: Age and composition effects. *Science* **171**, 567–571 (1971).
5. Moroz, L. V., Fisenko, A. V., Semjonova, L. F., Pieters, C. M. & Korotaeva, N. N. Optical effects of regolith processes on S-asteroids as simulated by laser shots on ordinary chondrite and other mafic materials. *Icarus* **122**, 366–382 (1996).
6. Sasaki, S., Nakamura, K., Hamabe, Y., Kurahashi, E. & Hiroi, T. Production of iron nanoparticles by laser irradiation in a simulation of lunar-like space weathering. *Nature* **410**, 555–557 (2001).
7. Chapman, C. R. S-type asteroids, ordinary chondrites, and space weathering: The evidence from Galileo's fly-bys of Gaspra and Ida. *Meteorit. Planet. Sci.* **31**, 699–725 (1996).
8. Clark, B. E. *et al.* Space weathering on Eros: Constraints from albedo and spectral measurements of Psyche crater. *Meteorit. Planet. Sci.* **36**, 1617–1637 (2001).
9. Ivezić, Ž., Jurić, M., Lupton, R. H., Tabachnik, S. & Quinn, T. in *Survey and Other Telescope Technologies and Discoveries* (eds Tyson, J. A. & Wolff, S.) Vol. 4836, 98–103 *Proc. SPIE* (2002).
10. Hapke, B. Space weathering from Mercury to the asteroid belt. *J. Geophys. Res.* **106**, 10039–10074 (2001).
11. Michel, P., Benz, W., Tanga, P. & Richardson, D. C. Collisions and gravitational reaccumulation: Forming asteroid families and satellites. *Science* **294**, 1696–1700 (2001).
12. Milani, A. & Knežević, Z. Asteroid proper elements and the dynamical structure of the asteroid main belt. *Icarus* **107**, 219–254 (1994).
13. Gaffey, M. J. *et al.* Mineralogical variations within the S-type asteroid class. *Icarus* **106**, 573–602 (1993).
14. Doressoundiram, A., Barucci, M. A., Fulchignoni, M. & Florczak, M. EOS family: A spectroscopic study. *Icarus* **131**, 15–31 (1998).
15. Deeming, T. J. Stellar spectral classification. *I. Mon. Not. R. Astron. Soc.* **127**, 493–516 (1964).
16. Ivezić, Ž. *et al.* Color confirmation of asteroid families. *Astron. J.* **124**, 2943–2948 (2002).
17. Tedesco, E. F., Noah, P. V., Noah, M. & Price, S. D. The supplemental IRAS minor planet survey. *Astron. J.* **123**, 1056–1085 (2002).
18. Sykes, M. V. *et al.* The 2MASS asteroid and comet survey. *Icarus* **146**, 161–175 (2000).
19. Bus, S. J. & Binzel, R. P. Phase II of the small main-belt asteroid spectroscopic survey. *Icarus* **158**, 106–145 (2002).
20. Migliorini, F. *et al.* Surface properties of (6) Hebe: A possible parent body of ordinary chondrites. *Icarus* **128**, 104–113 (1997).
21. Morbidelli, A. & Vokrouhlický, D. The Yarkovsky-driven origin of near-Earth asteroids. *Icarus* **163**, 120–134 (2003).
22. Marti, K. & Graf, T. Cosmic-ray exposure history of ordinary chondrites. *Annu. Rev. Earth Planet. Sci.* **20**, 221–243 (1992).
23. Binzel, R. P., Bus, S. J., Burbine, T. H. & Sunshine, J. M. Spectral properties of near-Earth asteroids: evidence for sources of ordinary chondrite meteorites. *Science* **273**, 946–948 (1996).
24. Rabinowitz, D. L. Size and orbit dependent trends in the reflectance colours of earth-approaching asteroids. *Icarus* **134**, 342–346 (1998).
25. di Martino, M., Manara, A. & Migliorini, F. 1993 VW: an ordinary chondrite-like near-Earth asteroid. *Astron. Astrophys.* **302**, 609–612 (1995).
26. Lazzarin, M., di Martino, M., Barucci, M. A., Doressoundiram, A. & Florczak, M. Compositional properties of near-Earth asteroids: spectroscopic comparison with ordinary chondrite meteorites. *Astron. Astrophys.* **327**, 388–391 (1997).
27. Gaffey, M. J. Spectral reflectance characteristics of the meteorite classes. *J. Geophys. Res.* **81**, 905–920 (1976).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank C. Chapman and J. Taylor for providing background on the S-complex conundrum and B. Clark for useful discussions and further studies. D.N. thanks NASA PG&G and the SwRI Quicklook programmes for providing support. Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the US Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS website is <http://www.sdss.org/>.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.J. (jedicke@ifa.hawaii.edu).

Tonks–Girardeau gas of ultracold atoms in an optical lattice

Belén Paredes¹, Artur Widera^{1,2,3}, Valentin Murg¹, Olaf Mandel^{1,2,3}, Simon Fölling^{1,2,3}, Ignacio Cirac¹, Gora V. Shlyapnikov⁴, Theodor W. Hänsch^{1,2} & Immanuel Bloch^{1,2,3}

¹Max-Planck-Institut für Quantenoptik, D-85748 Garching, Germany

²Sektion Physik, Ludwig-Maximilians-Universität, Schellingstrasse 4/III, D-80799 Munich, Germany

³Institut für Physik, Johannes Gutenberg-Universität, D-55099 Mainz, Germany

⁴Laboratoire Physique Théorique et Modèles Statistique, Université Paris Sud, Bâtiment 100, 91405 Orsay Cedex, France, and Van der Waals-Zeeman Institute, University of Amsterdam, 1018 XE Amsterdam, The Netherlands

Strongly correlated quantum systems are among the most intriguing and fundamental systems in physics. One such example is the Tonks–Girardeau gas^{1,2}, proposed about 40 years ago, but until now lacking experimental realization; in such a gas, the repulsive interactions between bosonic particles confined to one dimension dominate the physics of the system. In order to minimize their mutual repulsion, the bosons are prevented from occupying the same position in space. This mimics the Pauli exclusion principle for fermions, causing the bosonic particles to exhibit fermionic properties^{1,2}. However, such bosons do not exhibit completely ideal fermionic (or bosonic) quantum behaviour; for example, this is reflected in their characteristic momentum distribution³. Here we report the preparation of a Tonks–Girardeau gas of ultracold rubidium atoms held in a two-dimensional optical lattice formed by two orthogonal standing waves. The addition of a third, shallower lattice potential along the long axis of the quantum gases allows us to enter the Tonks–Girardeau regime by increasing the atoms' effective mass and thereby enhancing the role of interactions. We make a theoretical prediction of the momentum distribution based on an approach in which trapped bosons acquire fermionic properties, finding that it agrees closely with the measured distribution.

The physics of ultracold one-dimensional (1D) Bose systems is very different from that of ordinary three-dimensional (3D) cold gases^{1,2,4,5}. For example, by decreasing the particle density n , a usual 3D quantum many-body system becomes more ideal, whereas in a 1D Bose gas the role of interactions becomes more important. The reason is that at temperatures $T \rightarrow 0$, the kinetic energy of a particle at the mean interparticle separation is $K \propto n^2$ and it decreases with decreasing density n faster than the interaction energy per particle, $I \propto n$. The ratio of the interaction to kinetic energy, $\gamma = I/K$, characterizes the different physical regimes of the 1D quantum gas. For a large value of γ , the gas enters the Tonks–Girardeau (TG) regime, where the repulsion between particles strongly decreases the wavefunction at short interparticle distances.

Achieving such a TG regime and observing 'fermionization' of

the 1D Bose system is a great challenge, and it is complementary to the current experiments in which bosonic properties are observed in fermionic quantum gases^{6–9}. The 1D regime is obtained by tightly confining the particle motion in two directions to zero point oscillations^{4,5,10}. It was first demonstrated in experiments with weakly interacting Bose-condensed trapped gases, where $\gamma \ll 1$ (see refs 11, 12). In ref. 13, a tight radial confinement was realized by using two-dimensional (2D) optical lattice potentials to create an array of 1D quantum gases. In later experiments with optical lattices^{14,15} it has become possible to reach a 1D regime with $\gamma \approx 1$, that is, in between a weakly interacting 1D Bose condensed gas and a fermionized TG gas. So far, however, it has not been possible to bridge the last one or two orders of magnitude in γ that could bring the bosonic quantum gas fully into the TG regime. Larger values of γ could either be reached by decreasing the density of the quantum gas or by increasing the effective interaction strength between the particles^{4,5}.

In this work, we propose and demonstrate a novel way to achieve the TG regime. The main point is to include an additional optical lattice along the 1D gas, which results in an increase of γ . For a homogeneous gas, γ can be expressed as $\gamma = mg/\hbar^2 n$, where g is the 1D interaction strength, m the mass of a single atom, and $\hbar$ denotes Planck's constant divided by 2π . The addition of a periodic potential along the third axis increases the effective mass, and thus leads to an increase of γ . In fact, in the limit in which only the first Bloch band is occupied, we have $I = U\nu$ and $K = J\nu$, where ν is the filling factor, U the on-site interaction energy and J the tunnelling amplitude, and thus $\gamma = U/J$. Additionally, in order to achieve a pure TG regime in a lattice, the filling factor ν should be smaller than unity: otherwise, doubly occupied sites would be present, and the direct correspondence to the TG gas would be lost (as in a recent experiment, see ref. 16). Following these ideas, we have been able to enter the TG regime with $\gamma \approx 5$ –200. In this regime, the bosons can be theoretically described using a 'fermionization' approach^{17,18}.

For $\gamma \rightarrow \infty$, the ground state of N bosons at zero temperature is

described by the many-body wavefunction:

$$\Psi_0(x_1, x_2, \dots, x_N) \propto |\det[\varphi_i(x_j)]|, \quad i, j = 1, 2, \dots, N \quad (1)$$

where $\det$ denotes the Slater determinant, and $\varphi_i(x)$ is the i th eigenfunction of the single-particle hamiltonian. The presence of the Slater determinant guarantees that the wavefunction vanishes whenever two particles occupy the same position in space. However, the absolute value of the determinant ensures that the wavefunction for the bosons remains completely symmetric. This wavefunction reflects the fundamental similarities between strongly interacting bosons and non-interacting fermions in one dimension, with properties such as the spatial density distribution, the density–density correlation function, or the entropy of the gas being the same as in the case of non-interacting fermions. More interestingly though, several properties are strongly modified by the presence of the absolute value of the determinant, leading to a unique behaviour of, for example, the momentum distribution of the TG gas³. This can be understood qualitatively in the following way: the bosonic particles in a TG gas are not allowed to occupy the same position in space. Owing to this restriction, they are distributed over a more extended region in momentum space than in the case of an ideal or weakly interacting Bose gas. On the other hand, in order to keep themselves apart from each other, they do not need to be in different momentum states, as would be the case for fermions.

We first describe the experimental realization together with the measured data, and then provide a detailed theoretical analysis of the system. In order to reach the regime of low filling factor, we start with a rather small Bose–Einstein condensate (BEC) of approximately $(3\text{--}4) \times 10^4$ ⁸⁷Rb atoms in a magnetic trap. Then the BEC is loaded into a 2D optical lattice potential (along the y - and z -axes), such that an array of 1D quantum gases confined to narrow potential tubes is created (Fig. 1a). The lattice potential is formed by superimposing two orthogonal standing waves with a wavelength of 823 nm on top of the BEC. In order to transfer the atoms into the optical potential, the potential depth of the optical lattice is first gradually increased to a mean final value of $27 E_r$ (Fig. 1b). Here E_r is the recoil energy $\hbar^2 k^2/2m$, with k describing the wave vector of the lattice laser light. During this ramp up of the lattice potentials, the tunnel coupling between the different 1D quantum gases decreases exponentially. This results in a decoupling of the quantum gases, such that particle exchange between different tubes is suppressed. For the maximum lattice depth, the gaussian shape of the laser beams (160 μm waist) leads to an axial harmonic confinement of the 1D gases with a trapping frequency of $\omega_{ax} \approx 2\pi \times 60$ Hz. This has been verified by exciting a 'sloshing' motion of the thermal cloud and by parametric heating measurements, which both agree with the calculated value. Furthermore, the depths of all standing-wave potentials have been measured by vibrational band spectroscopy¹⁹. For such 1D quantum gases, without a lattice in the axial direction, we have $\gamma \approx 0.5$ near the lattice centre.

After a further hold time of 10 ms, we add an optical standing wave along the axial direction (x axis) in order to increase γ . The intensity of the laser forming this lattice potential (operated at a wavelength of 854 nm) is ramped up to a final depth V_{ax} of up to $18.5 E_r$. The axial momentum distribution of the quantum gases is subsequently probed by suddenly removing all optical and magnetic trapping potentials, and by imaging the atom clouds after a time-of-flight period of 16 ms. In order to prevent a strong expansion of the atom cloud along the propagation axis of the imaging laser beam (z axis), which would make the experiment more sensitive to misalignments in the imaging axis, we reduce the confinement along this axis by lowering the z -lattice potential to $6 E_r$ within a time of 100 μs before initiating the ballistic expansion sequence. Also, along the x axis we use a ramp down, which is not fully non-adiabatic and leads to a narrowing of the gaussian envelope in the observed momentum distribution by $\sim 20\%$. This enhances the number of atoms in the central momentum peak. From the

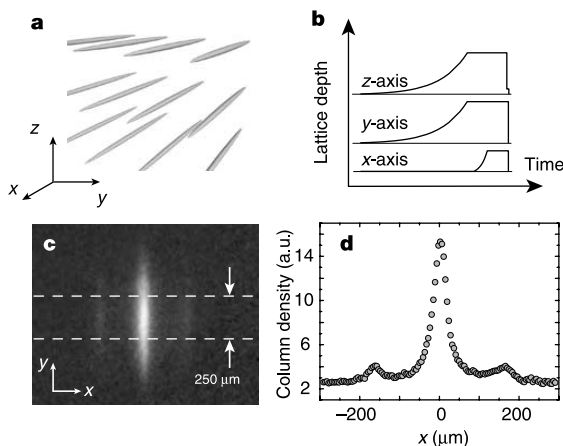


Figure 1 Experimental sequence and momentum profiles. **a**, Using a 2D optical lattice potential, we realize an array of 1D quantum gases. **b**, These quantum gases are created by first increasing the optical lattice depths along the y and z axes in an exponential ramp over a time of 160 ms (time constant $t = 40$ ms) to a mean final value of $27 E_r$. After a further hold time of 10 ms at this final lattice depth, we increase the optical lattice potential along the x axis within a time of 20 ms (time constant $t = 10$ ms) to a variable lattice depth V_{ax} . The quantum gases are then allowed to equilibrate for another 30 ms before we probe the momentum distribution as described in the text. **c**, Typical time-of-flight images after a ballistic expansion of the atom clouds over a time of 16 ms for an axial optical lattice depth $V_{ax} = 6.5 E_r$. The white dashed lines denote the area from which averaged momentum profiles along the x axis are extracted (**d**).

absorption images, we extract profiles of the axial momentum distribution by averaging horizontal profiles through the centre of the atom cloud (Fig. 1c).

In Fig. 2, we show six experimentally measured momentum profiles (see Supplementary Information for all 12 momentum profiles), corresponding to different values of the axial optical lattice depth ($V_{\text{ax}}/E_r = 0-18.5$). In Fig. 2a there is no lattice present along the x axis, and thus no first-order diffraction peak appears. Here, the value of γ is ~ 0.5 at the trap centre. For the rest of the figures (Fig. 2b–f) we can use the relation $\gamma \approx U/J$ obtaining $\gamma \approx 5-200$, which indicates that the TG regime is entered rather

rapidly when increasing the axial lattice depth. In Fig. 2b–f we also plot our theoretical predictions based on fermionization at finite temperature averaged over the different 1D tubes (see Methods). Apart from a normalizing factor for each experimental curve, only the atom number in the central tube is used as an overall adjustable parameter in this model. This atom number is, however, kept constant between different momentum profiles. The initial temperature for the lowest axial lattice depth $V_{\text{ax}} = 4.6 E_r$ has been obtained through a finite temperature fit to the corresponding momentum profiles using our fermionization approach. From this initial temperature, the temperatures of the quantum gases at

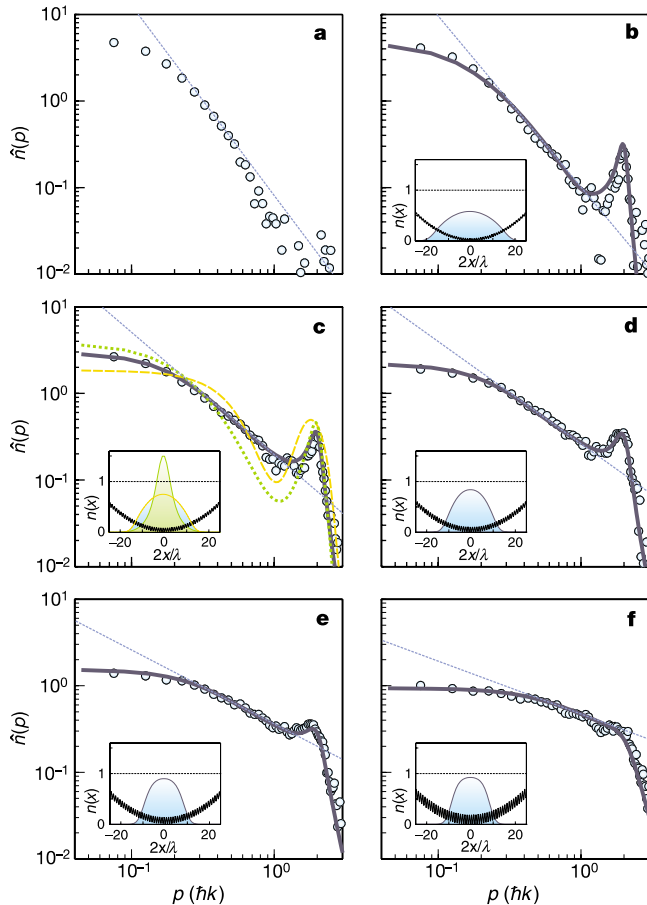


Figure 2 Momentum profiles of the 1D quantum gases for different axial lattice depths. In **b–f**, the experimental data (blue circles) are displayed together with our theoretical predictions (black line) based on fermionization at finite temperatures, averaged over the different 1D tubes. In order to emphasize the linear part of the momentum profiles, an auxiliary straight line with the corresponding slope is shown in each plot. In **c**, the momentum profiles for the ideal Bose gas (green dotted lines) and the ideal Fermi gas (yellow dashed lines) are also displayed for comparison. For all plots, an atomic distribution characterized by an atom number $N_{0,0} = 18$ in the central tube is used, for which we have found the best agreement with the experimental data (see Methods). In the insets of **b–f**, the density profile of a single 1D tube with $N = 15$ particles at the corresponding temperature and lattice depth is shown for the fermionized gas (black lines in plots **b–f**), for the ideal Fermi gas (yellow line in **c**), and for the ideal Bose gas (green line in **c**). The values of the axial lattice depths V_{ax} , the average temperatures, the slopes α of the linear part of the momentum profiles, and the values of $\gamma = U/J$ are: **b**, $4.6 E_r$ and $k_B T/J = 0.5$ (Tonks), $\alpha = 1.90$, $\gamma = 5.5$; **c**, $7.4 E_r$ and $k_B T/J = 0.7$ (Tonks), $k_B T/J = 1.6$ (ideal Bose gas), $k_B T/J = 0.7$ (ideal Fermi gas), $\alpha = 1.4$, $\gamma = 13.7$; **d**, $9.3 E_r$ and $k_B T/J = 0.9$ (Tonks), $\alpha = 1.2$, $\gamma = 23.6$; **e**, $12 E_r$ and $k_B T/J = 1.3$ (Tonks), $\alpha = 0.8$, $\gamma = 47.6$; and **f**, $18.5 E_r$ and $k_B T/J = 3.9$ (Tonks), $\alpha = 0.6$, $\gamma = 204.5$. For the momentum profile without the axial lattice (**a**), we find $\alpha = 2.2$ and $\gamma = 0.5$ at the centre of the trap.

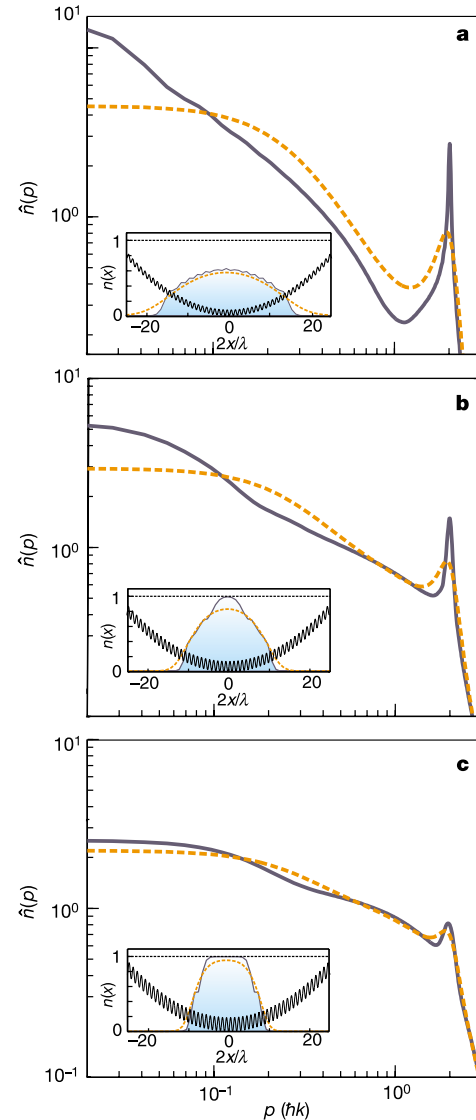


Figure 3 Momentum profiles of a single 1D tube obtained from our fermionization-based theory for different lattice depths. The plots are shown for axial lattice depths V_{ax} of $5.0 E_r$ (**a**), $9.5 E_r$ (**b**), and $12.0 E_r$ (**c**). For all plots, the number of particles is $N = 15$, and $b = 8 \times 10^{-4} E_r$ (this value of b corresponds to the trapping frequency of the experiment; see Methods). In each plot, the log-log momentum profile at $k_B T/J = 0$ (black line) is displayed together with that at $k_B T/J = 1.0$ (orange dashed line). The density profiles at $k_B T/J = 0$ and $k_B T/J = 1.0$, together with the corresponding lattice-harmonic potential, are shown in the inset of each plot. Note that at $k_B T/J = 0$, finite size effects make the slope at low momenta deviate from the ideal $1/2$. The slope α is larger than $1/2$ for small filling factors ($\alpha = 0.79$ in **a**), it approaches $1/2$ as a Mott phase is developed at the centre of the trap ($\alpha = 0.49$ in **b**), and it decreases to zero deep in the Mott phase ($\alpha = 0.29$ in **c**).

increasing lattice depth V_{ax} have been modelled by assuming conservation of entropy during the ramp up of the axial lattice. For all 12 experimentally measured momentum profiles (see Supplementary Information), we find excellent agreement with the theory based on fermionization. For reference, we have plotted the results obtained assuming an ideal Bose or Fermi gas, also averaged over all the 1D tubes and at finite temperatures (see for example, Fig. 2c).

The observed momentum distributions allow us to conclude that we are dealing with a finite, non-uniform, TG gas in a lattice. Our results show pronounced deviations²⁰ from the generic behaviour of the uniform TG gas at zero temperature, where one has a $1/p^{1/2}$ low momentum distribution³ giving a slope of 1/2 in the log-log plot. In our experiment we observe: (1) a rather flat momentum distribution at small momenta p , and (2) a linear region at larger p , with a slope decreasing with an increase in the lattice depth. This behaviour is in excellent agreement with the predictions of our fermionization-based theory for such a TG gas with a finite number of particles (about 20 per tube) in a lattice, in the presence of a harmonic trap, and at finite temperatures (see Fig. 3). Note that already for our lowest axial lattice depths we find $\gamma \gg 1$, and a further increase in the lattice depth mainly changes the average filling factor in our system and allows us to study the behaviour of a TG gas at different densities.

Most of the relevant physics concerning the momentum distribution for our case can be qualitatively understood by considering a uniform lattice system at the same temperature and with a filling factor equal to the average filling factor of the trapped system. We can then restrict the discussion to the case $\nu \leq 1/2$, because for $\nu > 1/2$ the system can be viewed as a system of holes at filling factor $1 - \nu$. The filling factor determines a characteristic momentum $p_\nu = \hbar \times 2\pi\nu/\lambda$ related to the mean interparticle separation, where λ is the wavelength of the lattice laser light. At zero temperature, for $p \ll p_\nu$, the momentum distribution should exhibit a linear 1/2 behaviour, whereas for larger momenta short-range correlations²¹ tend to increase the slope. An increase in the filling factor, and therefore a decrease in the average separation between particles, modifies p_ν and can therefore lead to a change of the observed slope. Note that for the case of $\nu = 1/2$, the momentum p_ν is the closest to the lattice momentum $\hbar \times 2\pi/\lambda$, and the momentum distribution is the least affected by short-range correlations. At finite temperatures a new momentum scale sets in^{18,22}, below which the slope has a tendency to decrease. This is the momentum $p_T = \hbar \times \pi/L_T$, where $L_T \approx \lambda/k_B T \sin \pi\nu$ is a characteristic length of thermal fluctuations. For a small filling factor, this length coincides with the gas phase result $L_T \approx \hbar^2 n/m^* k_B T$, in which the particle mass is replaced by the effective mass $m^* = 2\hbar^2/J\lambda^2$. In our experiment we have $p_T \approx p_\nu$. Therefore, finite-temperature effects overlap with effects of short-range correlations and we observe a rather flat momentum distribution at small p , and a linear region with slope larger than 1/2 for larger p .

The presence of the harmonic trapping potential introduces important changes in the observed momentum profiles. First, in contrast to the uniform case, an adiabatic increase of the lattice depth increases the ratio $k_B T/J$. This increases the momentum p_T and the flat region extends to larger momenta. Second, the slope of the linear part decreases with the lattice depth, and the generic 1/2 value is recovered on approach to the Mott insulator transition^{16,23–26}. This is a fundamental feature that is present irrespective of the number of particles and trap frequency. It is related to the fact that in the trapped case the characteristic average filling factor of the system increases with the lattice depth, because the tunnelling amplitude J decreases and particles try to accumulate near the trap centre. At the Mott insulator crossover, where the filling factor at the trap centre is equal to unity, the average filling factor is close to $\nu = 1/2$ (see Methods). This is the value for which the effects of short-range correlations are strongly suppressed in the homo-

geneous lattice system, and one comes closest to the generic behaviour with slope 1/2. Last, we note that in the weakly interacting regime for a trapped quasicondensate, one should have a lorentzian momentum distribution²⁷, which would give a slope close to 2 for $p \gg \hbar \times \pi/L_T$. Already, for low axial lattice depths V_{ax} we observe a smaller slope, which emphasizes a strong difference of our system from previously studied 1D quasicondensates.

In summary, we have prepared a TG gas in an optical lattice. Here the bosonic atoms exhibit a pronounced fermionic behaviour, and show a momentum distribution that is in excellent agreement with a theory of fermionized trapped Bose gases. In a next step, it will be intriguing to use photoassociation in optical lattices to probe the reduced two-body correlations, which are expected in a TG gas²⁸. Furthermore, by using two bosonic atomic species and tuning the sign and strength of the atomic interactions, it should be possible to observe a behaviour similar to strongly correlated fermions. For example, the bosonic atoms can undergo a BCS transition and form Cooper pairs in the same way as electrons do in a superconductor²⁹. $\square$

Methods

Description of the 1D quantum gases using fermionization

Here we develop the theoretical treatment based on fermionization that we have used above to model the experiment. We consider N bosonic atoms moving in the lowest band of a 1D lattice and experiencing an additional harmonic potential. This situation is described by the Bose–Hubbard hamiltonian $H = H_B + V$, where

$$H_B = -J \sum_{\ell=-\infty}^{\infty} (a_{\ell}^{\dagger} a_{\ell+1} + a_{\ell+1}^{\dagger} a_{\ell}) + b \sum_{\ell=-\infty}^{\infty} \ell^2 a_{\ell}^{\dagger} a_{\ell}$$

$$V = U \sum_{\ell=-\infty}^{\infty} a_{\ell}^{\dagger 2} a_{\ell}^2$$

The term H_B describes the motion of the atoms in the combined lattice-harmonic potential, and the term V accounts for on-site interactions. The bosonic operators a_{ℓ} annihilate one boson at the ℓ th site, and fulfil canonical commutation relations $[a_{\ell}, a_{\ell'}^{\dagger}] = \delta_{\ell, \ell'}$. The parameter b is related to the frequency ω of the harmonic potential by $b = 1/8 m \omega^2 \lambda^2$.

We are interested in the strongly interacting or Tonks regime, in which two atoms cannot occupy the same lattice site. Within this regime, the bosonic operators a_{ℓ} can be re-expressed using the Jordan–Wigner transformation³⁰ (JWT) in terms of fermionic ones c_{ℓ} fulfilling $[c_{\ell}, c_{\ell'}^{\dagger}]_{+} = \delta_{\ell, \ell'}$. Under the JWT, the interacting Bose hamiltonian H_B is transformed into a non-interacting fermionic hamiltonian H_F through the replacement $a_{\ell} \rightarrow c_{\ell}$. In order to predict the behaviour of the different bosonic observables, one has to transform them into fermionic ones via the JWT, and then evaluate the corresponding expectation values for the fermionic ground state. At $T = 0$, the fermionic ground state is given by the Slater determinant of equation (1). At a finite temperature T , the wavefunction is a mixture of different Slater determinants characterized by the many-body density matrix $\rho \propto \exp(-H_F/k_B T)$, where k_B is Boltzmann's constant.

Density and momentum distributions of fermionized Bose gases

The particle density $n(x)$ coincides with that of non-interacting fermions, as the JWT maps the corresponding bosonic observable onto the same fermionic one (that is $a_{\ell}^{\dagger} a_{\ell} \rightarrow c_{\ell}^{\dagger} c_{\ell}$). Under the Thomas–Fermi approximation we have:

$$n(x) = \frac{1}{\pi} \arccos \left(\max \left[\frac{\mu - bx^2}{-2J}, -1 \right] \right)$$

if $\mu - bx^2 > -2J$ and zero otherwise. The size L of the cloud is $L = \lambda \sqrt{(2J + \mu)/4b}$, and μ is determined by imposing the condition that the total number of particles is N . When $\mu \geq 2J$ a Mott phase is produced at the centre of the trap, and $n(x=0)$ is equal to 1. At this point the average filling factor of the system $\bar{\nu} = \lambda N/2L \approx 3/\sqrt{2}\pi$, a value which is close to 1/2.

The momentum distribution $\hat{n}(p)$ is related to the one-particle correlation function $\langle a_{\ell}^{\dagger} a_{\ell'} \rangle$ through:

$$\hat{n}(p) = |\Phi(p)|^2 \sum_{\ell, \ell'=-\infty}^{\infty} e^{-ip(\ell-\ell')} \langle a_{\ell}^{\dagger} a_{\ell'} \rangle$$

where $\Phi(p)$ is the Fourier transform of the Wannier function, and p denotes momentum in units of $\hbar k$. Using the JWT, the bosonic one-particle correlation function can be re-expressed as:

$$\langle a_{\ell}^{\dagger} a_{\ell'} \rangle = \langle c_{\ell}^{\dagger} (-1)^{\sum_{m=\ell'-1}^{\ell-1} c_m^{\dagger} c_m} c_{\ell'} \rangle, \ell > \ell'$$

Making extensive use of Wick's theorem, one can re-express this quantity as a Töplitz determinant $\langle a_{\ell}^{\dagger} a_{\ell'} \rangle = \det[G_{\ell, \ell'}]$, where $G_{\ell, \ell'}$ is a $\ell - \ell' + 1$, $\ell - \ell' - 1$ matrix with elements $(G_{\ell, \ell'})_{x, y} = \langle c_{\ell'+y-1}^{\dagger} c_{\ell+x} \rangle - \delta_{x, y-1}/2$.

Therefore, in order to evaluate the momentum distribution at a finite temperature T one has to determine the one-particle correlation functions for a non-interacting Fermi

system at that temperature. We have used the grand canonical Fermi–Dirac distribution and the exact eigenstates $\varphi_i(x)$ of the single-particle hamiltonian to determine the momentum distribution in this way.

Averaging over the array of 1D quantum gases

In order to give a quantitative prediction for the experimental situation, we have averaged the momentum distribution for different tubes. To determine the atomic distribution, we have assumed that during the ramp up of the 2D optical lattice potential, tunnelling becomes negligible, and we have an array of independent 1D gases. For each tube, we have assumed a Thomas–Fermi density profile. Minimizing the total energy of the array with respect to the number of atoms in each of the tubes, we obtain $N_{ij} = N_{0,0} \left(1 - \frac{5N}{2\pi N_{0,0}} (i^2 + j^2)\right)^{3/2}$, where N_{ij} is the number of atoms in a tube located at position (i, j) in the 2D optical lattice, N is the total number of particles in the array, and $N_{0,0}$ is the number of particles in the central tube. It follows that the probability of having a tube with M particles is:

$$P(M) = \frac{2}{3} \frac{1}{N_{0,0}^{2/3} M^{1/3}}, \quad M \leq N_{0,0}.$$

Remarkably, this distribution only depends on one parameter, namely, the number of particles in the central tube, which is the only adjustable parameter in our model.

The temperature of each 1D quantum gas has been calculated assuming adiabatic evolution of the system during the ramp up of the axial lattice. Owing to the presence of the harmonic confinement, the ratio $k_B T/J$ is not conserved in the adiabatic evolution. Given the temperature at $V_{ax} = 4.6 E_T$ (see Supplementary Information), the conservation of entropy allows us to determine the temperature at the final lattice depth V_{ax} . The entropy of the TG gas coincides with that of the non-interacting Fermi gas, as both have the same spectrum and density of states. This results in the same temperatures for a TG gas and an ideal Fermi gas, but a different temperature for the ideal Bose gas when the axial lattice depth is increased. Note that tubes with different number of particles also have different temperatures at the same lattice depth.

Received 21 February; accepted 1 April 2004; doi:10.1038/nature02530.

- Girardeau, M. Relationship between systems of impenetrable bosons and fermions in one dimension. *J. Math. Phys.* **1**, 516–523 (1960).
- Lieb, E. H. & Liniger, W. Exact analysis of an interacting Bose gas. The general solution and the ground state. *Phys. Rev.* **130**, 1605–1616 (1963).
- Lenard, A. Momentum distribution in the ground state of the one-dimensional system of impenetrable bosons. *J. Math. Phys.* **5**, 930–943 (1964).
- Petrov, D. S., Shlyapnikov, G. V. & Walraven, J. T. M. Regimes of quantum degeneracy in trapped 1D gases. *Phys. Rev. Lett.* **85**, 3745–3749 (2000).
- Dunjko, V., Lorent, V. & Olshanii, M. Bosons in cigar-shaped traps: Thomas-Fermi regime, Tonks-Girardeau regime, and in between. *Phys. Rev. Lett.* **86**, 5413–5416 (2001).
- Jochim, S. *et al.* Bose-Einstein condensation of molecules. *Science* **302**, 2101–2103 (2003).
- Greiner, M., Regal, C. & Jin, D. S. Emergence of a molecular Bose-Einstein condensate from a Fermi gas. *Nature* **426**, 537–540 (2003).
- Zwierlein, M. W. *et al.* Observation of Bose-Einstein condensation of molecules. *Phys. Rev. Lett.* **91**, 250401 (2003).
- Regal, C., Greiner, M. & Jin, D. S. Observation of resonance condensation of fermionic atom pairs. *Phys. Rev. Lett.* **92**, 040403 (2004).
- Olshanii, M. Atomic scattering in the presence of an external confinement. *Phys. Rev. Lett.* **81**, 938–941 (1998).
- Goerlitz, A. *et al.* Realization of Bose-Einstein condensates in lower dimensions. *Phys. Rev. Lett.* **87**, 130402 (2001).
- Schreck, F. *et al.* A quasipure Bose-Einstein condensate immersed in a Fermi sea. *Phys. Rev. Lett.* **87**, 080403 (2001).
- Greiner, M., Bloch, I., Mandel, O., Hänsch, T. W. & Esslinger, T. Exploring phase coherence in a 2D lattice of Bose-Einstein condensates. *Phys. Rev. Lett.* **87**, 160405 (2001).
- Moritz, H., Stöferle, T., Köhl, M. & Esslinger, T. Exciting collective oscillations in a trapped 1D gas. *Phys. Rev. Lett.* **91**, 250402 (2003).
- Laburthe Tolra, B. *et al.* Observation of reduced three-body recombination in a fermionized 1D Bose gas. Preprint at (<http://xxx.lanl.gov/cond-mat/0312003>) (2003).
- Stöferle, T., Moritz, H., Schori, C., Köhl, M. & Esslinger, T. Transition from a strongly interacting 1D superfluid to a Mott insulator. *Phys. Rev. Lett.* **92**, 130403 (2004).
- Eftov, K. B. & Larkin, A. I. Correlation functions in one-dimensional systems with strong interactions. *Sov. Phys. JETP* **42**, 390–396 (1976).
- Korepin, V. E., Bogoliubov, N. M. & Izergin, A. G. *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge Univ. Press, Cambridge, 1993).
- Ovchinnikov, Y. B. *et al.* Diffraction of a released Bose-Einstein condensate by a pulsed standing light wave. *Phys. Rev. Lett.* **83**, 284–287 (1999).
- Astrakharchik, G. E. & Giorgini, S. Correlation functions and momentum distributions of one-dimensional Bose systems. *Phys. Rev. A* **68**, 031602 (2003).
- Olshanii, M. & Dunjko, V. Short-distance correlation properties of the Lieb-Liniger system and momentum distributions of trapped one-dimensional atomic gases. *Phys. Rev. Lett.* **91**, 090401 (2003).
- Cazalilla, M. A. Bosonizing one-dimensional cold atomic gases. *J. Phys. B* **37**, S1–S47 (2004).
- Fisher, M. P. A., Weichman, P. B., Grinstein, G. & Fisher, D. S. Boson localization and the superfluid-insulator transition. *Phys. Rev. B* **40**, 546–570 (1989).
- Jaksch, D., Bruder, C., Cirac, J. I., Gardiner, C. W. & Zoller, P. Cold bosonic atoms in optical lattices. *Phys. Rev. Lett.* **81**, 3108–3111 (1998).
- Greiner, M., Mandel, O., Esslinger, T., Hänsch, T. W. & Bloch, I. Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms. *Nature* **415**, 39–44 (2002).
- Kollath, C., Schollwöck, U., von Delft, J. & Zwerger, W. Spatial correlations of trapped one-dimensional bosons in an optical lattice. *Phys. Rev. A* **69**, 031601 (2004).
- Richard, S. *et al.* Momentum spectroscopy of 1D phase fluctuations in Bose-Einstein condensates. *Phys. Rev. Lett.* **91**, 010405 (2003).

- Gangardt, D. M. & Shlyapnikov, G. V. Stability and phase coherence of trapped 1D Bose gases. *Phys. Rev. Lett.* **90**, 010401 (2003).
- Paredes, B. & Cirac, J. I. From Cooper pairs to Luttinger liquids with bosonic atoms in optical lattices. *Phys. Rev. Lett.* **90**, 150402 (2003).
- Sachdev, S. *Quantum Phase Transitions* (Cambridge Univ. Press, Cambridge, 1999).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank F. Gerbier, D. Gangardt and M. Olshanii for discussions, and M. Greiner for help in setting up the experiment. I.B. also acknowledges support from AFOSR.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to I.B. (bloch@uni-mainz.de).

Synthesis and characterization of chiral mesoporous silica

Shunai Che¹, Zheng Liu^{2,3}, Tetsu Ohsuna³, Kazutami Sakamoto⁴, Osamu Terasaki³ & Takashi Tatsumi⁵

¹Department of Chemistry, School of Chemistry and Chemical Technology, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai, 200240, China

²Bussan Nanotech Research Institute, 2-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

³Structural Chemistry, Arrhenius Laboratory, Stockholm University, S-10691 Stockholm, Sweden

⁴AminoScience Laboratory, Ajinomoto Co., Inc., 1-1 Suzuki-cho, Kawasaki 210-8681, Japan

⁵CREST, JST, Division of Materials Science and Chemical Engineering, Faculty of Engineering, Yokohama National University, 79-5 Tokiwadai, Yokohama 240-8501, Japan

Chirality is widely expressed in organic materials, perhaps most notably in biological molecules such as DNA, and in proteins, owing to the homochirality of their components (D-sugars and L-amino acids). But the occurrence of large-scale chiral pores in inorganic materials is rare¹. Although some progress has been made in strategies to synthesize helical and chiral zeolite-like materials^{1–3}, the synthesis of enantiomerically pure mesoporous materials is a challenge that remains unsolved⁴. Here we report the surfactant-templated synthesis of ordered chiral mesoporous silica, together with a general approach for the structural analysis of chiral mesoporous crystals by electron microscopy. The material that we have synthesized has a twisted hexagonal rod-like morphology, with diameter 130–180 nm and length 1–6 μm . Transmission electron microscopy combined with computer simulations confirm the presence of hexagonally ordered chiral channels of 2.2 nm diameter winding around the central axis of the rods. Our findings could lead to new uses for mesoporous silica and other chiral pore materials in, for example, catalysis and separation media, where both shape selectivity and enantioselectivity⁵ can be applied to the manufacturing of enantiomerically pure chemicals and pharmaceuticals.

We recently discovered a templating route for preparing well-ordered mesoporous silicas based on the self-assembly of chiral anionic surfactants and inorganic precursors by using aminosilane or quaternized aminosilane as a co-structure-directing agent (CSDA)⁶, which provided a potential method to synthesize mesoporous materials with inherent chirality. Among the anionic surfactants tested in our previous work, N-acyl-L-alanine is a chiral organic molecule that can form a chiral nematic phase in the presence of small amounts of decanol^{7,8}. This phenomenon has

been explained by the change in effective cross-sectional area at the aggregate interface⁷.

Our approach to the synthesis of chiral mesoporous silica is to control the micelle head group interaction between the following: (1) chiral anionic surfactant sodium N-acyl-L-alanate, (2) a small amount of the corresponding free amino acid N-acyl-L-alanine, and (3) quaternized aminosilane or aminosilane added as CSDA. We have achieved the formation of a chiral mesostructure by using N-miristoyl-L-alanine sodium salt (C_{14} -L-AlaS) partially neutralized with HCl, and N-trimethoxysilylpropyl-N,N,N-trimethylammonium chloride (the cation denoted here as TMAPS) or 3-aminopropyltrimethoxysilane (APS). As shown in Fig. 1, the three components self-assemble to form a micelle. The free form of N-miristoyl-L-alanine (denoted as C_{14} -L-AlaA) acts as a co-surfactant in a similar manner to decanol in the chiral liquid-crystal system^{7,8}. The presence of C_{14} -L-AlaA induces changes in the conformation of the amphiphile in the micelle and distorts the micellar shape, leading to chiral micelles^{7,8}. The positively charged ammonium site of TMAPS or APS interacts electrostatically with the negatively charged head group of the C_{14} -L-AlaS or C_{14} -L-AlaA, through double decomposition of C_{14} -L-AlaS and TMAPS or the neutralization of the C_{14} -L-AlaA and APS, respectively. The alkoxy-silane sites of TMAPS and APS are co-condensed with tetraethoxysilane (TEOS), to be assembled subsequently to form the silica

framework. The trimethylene groups of TMAPS and APS covalently tether the silicon atoms incorporated into the framework to the cationic ammonium groups. By adjusting the HCl/ C_{14} -L-AlaS and CSDA/ C_{14} -L-AlaS molar ratios, a pH and composition favourable for the formation of a chiral mesostructure can be achieved.

Figure 2 shows the X-ray diffraction (XRD) pattern of a calcined chiral mesoporous silica. Three well-resolved peaks in the range $2\theta = 1.5$ – 6° were indexed as 10, 11 and 20 reflections, based on the two-dimensional hexagonal $p6mm$ unit cell (which was also confirmed by transmission electron microscopy, TEM) with $a = 4.4$ nm, indicating that this mesoporous silica has a highly ordered hexagonal mesoscopic structure. Nitrogen adsorption-desorption isotherms also confirmed the existence of uniform mesopores with a BJH (Barrett-Joyner-Halenda) pore diameter of 2.2 nm. The BET (Brunauer-Emmett-Teller) surface area and mesopore volume were $600 \text{ m}^2 \text{ g}^{-1}$ and $370 \text{ mm}^3 \text{ g}^{-1}$, respectively (Supplementary Fig. 1).

Scanning electron microscopy (SEM) reveals that this material has a twisted hexagonal rod-like morphology (with a hexagonal

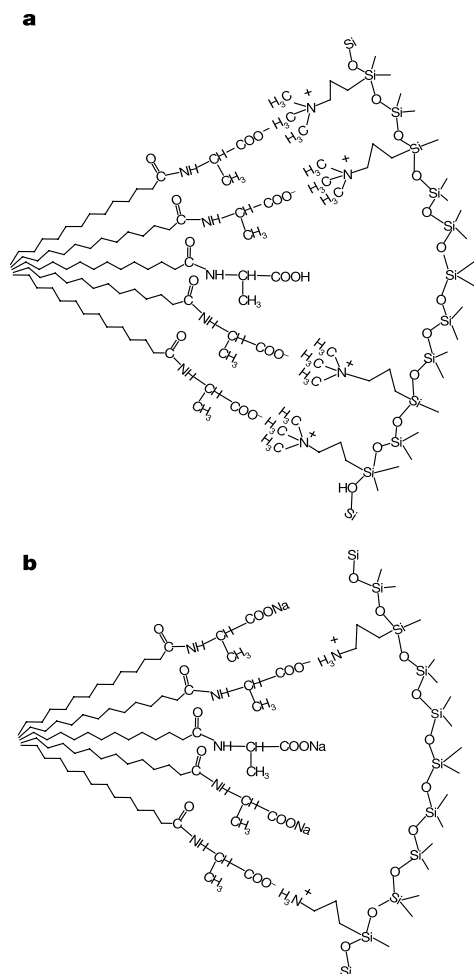


Figure 1 Schematic illustration of the two types of interaction between the head group of C_{14} -L-AlaS, its free amino acid C_{14} -L-AlaA and amino groups. The two types of interactions are achieved through double decomposition of negatively charged C_{14} -L-AlaS with positively charged quaternized aminosilane TMAPS (a), and neutralization of acid C_{14} -L-AlaA with primary aminosilane APS (b).

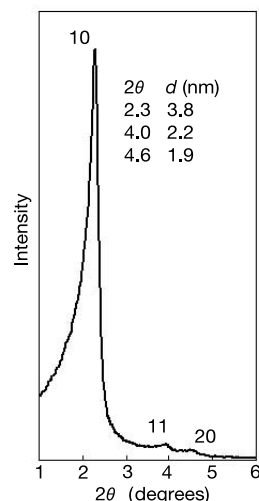


Figure 2 XRD pattern of calcined chiral mesoporous silica, indicating the hexagonal order. The chemical composition (mol ratios) of the reaction mixture was C_{14} -L-AlaS:HCl:TMAPS:TEOS: H_2O 1:0.16:0.43:7:1,722. XRD patterns were recorded on an MX Labo powder diffractometer using Cu $K\alpha$ radiation (40 kV, 20 mA), at the rate of $1.0^\circ \text{ min}^{-1}$ over the range $2\theta = 1.5$ – 10.0° .

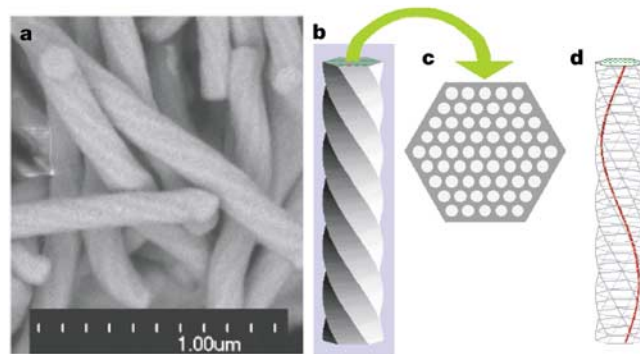


Figure 3 SEM image and schematic drawings of a structural model of chiral mesoporous silica. **a**, SEM image, showing the microscopic features of this sample (Hitachi S-5200 SEM). The samples were observed without any metal coating. **b**, Schematic drawing of a structural model of the chiral mesoporous material for TEM image simulation. **c**, Cross-section. **d**, One of the chiral channels in the material.

cross-section), 130–180 nm in outer diameter and $\sim 1\text{--}6\text{ }\mu\text{m}$ in length. From six distinct surfaces, the helical pitch along the rod axis was estimated to be $\sim 1.5\text{ }\mu\text{m}$ (Fig. 3a). We note that the external morphology of twisted hexagonal rods does not confirm the existence of a chiral pore structure inside the rods. However, the TEM images may be simulated using such a structural model of twisted, hexagonally shaped rods with two-dimensional chiral channels running inside, parallel to the ridge edge (Fig. 3b–d).

Three-dimensional structural information about an object can be obtained directly from a two-dimensional projection image (obtained by TEM) via either electron crystallography or tomographic techniques. However, here we used TEM images and specially developed computer simulation software to investigate the mesoporous structure of this chiral material. Figure 4 shows the TEM image of the rod (Fig. 4a), an enlarged image of the rectangular area shown in Fig. 4a (Fig. 4b), further enlargement of the rectangular part of Fig. 4b (Fig. 4c), and the corresponding simulated image (Fig. 4d). It can be seen that the simulated image

fits the TEM image very well. Although it is difficult to determine unambiguously whether the channels are chiral directly from the TEM images, the simulated image shows that these two-dimensional-hexagonal chiral channels running inside the rod are indeed consistent with the data. We note that there exist two kinds of fringes in the TEM images, indicated by arrows and arrowheads, which correspond to the interplanar spacing (10) and (11), respectively. When one of the $\langle 01 \rangle$ directions is parallel to the electron beam, the (10) interplanar spacing fringe appears. Consequently, between two sets of (10) fringes, the rod twists by 60° , which means that the distance between two sets of (10) fringes is one-sixth of one pitch (in the TEM image, this corresponds to the distance between two neighbouring arrows). The (11) fringes appear exactly in the middle between two sets of (10) fringes. The dark rhombic shape contrast in the simulated image is caused by the projections of $\langle 01 \rangle$ edges: this is a thickness effect due to the material having a hexagonal rod shape. Such contrast is appreciable in the TEM images.

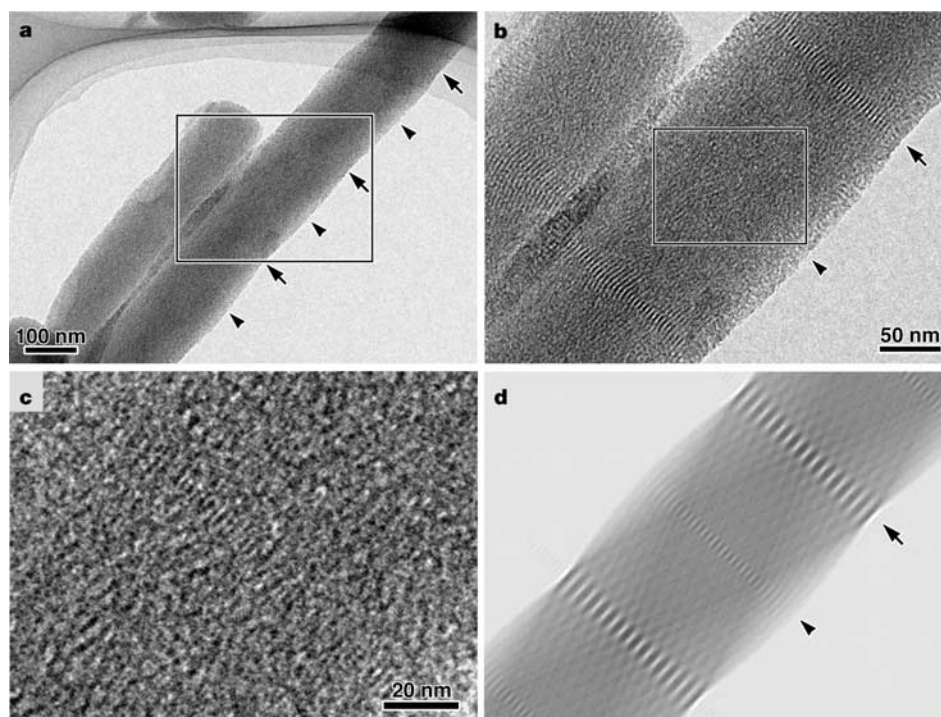


Figure 4 TEM images of chiral mesoporous silica. **a–c**, TEM images with different enlargements, showing two types of fringes (indicated by arrows and arrowheads) with different spacings. **d**, A simulated TEM image, showing good correspondence with the

observed image. Owing to the limits on the calculation time, computer memory and image resolution, the rod diameter and the number of channels inside the rod are reduced from those observed. Images were obtained using a JEOL JEM-3010 TEM at 300 kV.

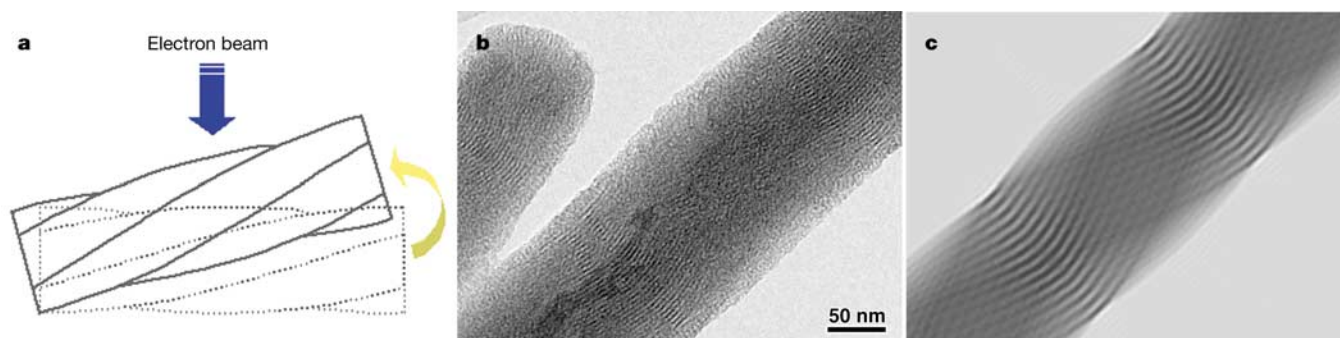


Figure 5 Schematic drawing of chiral mesoporous silica, its TEM image and the computer-simulated TEM image. The schematic drawing (**a**) shows the condition of the incident beam direction relative to the rod direction. The curvature of the (10) fringes can

be observed in both the TEM image (**b**) and the simulated image (**c**). In this simulation, the angle between the incident beam direction and the perpendicular direction of the rod was 15° .

We also note that some (10) fringes are longer and more curved. This is because, in such places, the rod is not precisely perpendicular to the electron beam. The variation of the (10) fringes at various places in one rod means that the rod is not perfectly straight. Figure 5 shows the effect of crystal tilting on the fringes in the TEM image: the image (Fig. 5b) can be closely simulated (Fig. 5c) when the tilting (Fig. 5a) is taken into consideration. It is important to note that all of the (10) fringes are distinctly curved and, furthermore, all the fringes curve in the same direction within one rod. It is confirmed by simulation that rods of right- and left-handedness have oppositely curved (10) fringes when they are tilted to the same direction relative to the electron beam. Therefore, the handedness of the rod may be determined by the direction of curvature of the (10) fringe. The rod observed in the TEM image shown has left-handed chirality.

This type of chiral rod, with mesoporous channels inside, can serve as an interesting container for fabricating chiral metal nanowires. In this work we succeeded in fabricating Co and Pt nanowires within the pores of the chiral mesoporous silica. TEM images (Supplementary Fig. 2) show Co and Pt wires (black contrast) inserted into the chiral channels, although there are also some large metal particles on the outer surface of the rod. Supplementary Fig. 2c shows a schematic drawing of channels with different curvature within which wires were fabricated. The two Co wires indicated by arrows and arrowheads clearly show that the curvatures are different. This is also a clear demonstration that the channels have chirality.

To characterize the enantiomeric purity of the material, CD (circular dichroism) spectra of the synthesis-gel were measured. These spectra showed a CD band around 210–250 nm with positive sign (Supplementary Fig. 3). This CD band is attributed to the chiral micelle, similar to acylglutamate⁹, which confirmed the occurrence of ordered, enantiopure self-assembly¹⁰. No CD band was detected for the synthesis-gel with racemic C₁₄-LD-AlaS (not shown). Handedness of the porous materials was estimated by counting characteristic morphologies from 500 randomly chosen crystals in the SEM images, and left-/right-handed ratios proved to be 6.5/3.5 and 7.5/2.5 for the C₁₄-L-AlaS/TMAPS and C₁₄-L-AlaA/APS synthesis systems, respectively. These findings suggest that competing driving forces for the formation of chiral structure exist, other than the chiral surfactant packing. As a control, a mixture of racemic sodium N-acyl-DL-alanate was also used, resulting in unstable two-dimensional-hexagonal mesoporous silica (the structure was destroyed by calcination) with only a small amount of twisted material formed. The control of enantiopurity and the synthesis mechanism is currently being studied.

To our knowledge, this is the first synthesis of silica crystals with chiral mesopores—which are not soft, but hard materials—by the structure-directing method. Such chiral materials have a clear advantage in their physical strength over chiral liquid-crystal assemblies, yielding a permanent chiral space. Natural chiral imprinting in inorganic structures probably occurs via a similar inorganic/organic chiral interface. □

Methods

Preparation of chiral mesoporous silica

Chiral mesoporous silica was synthesized by using chiral surfactant N-mristoyl-L-alanine sodium salt (C₁₄-L-AlaS) as template, and TMAPS and APS as CSDA. In a typical synthesis, C₁₄-L-AlaS (0.32 g, 1 mmol) was dissolved in deionized water (32 g) with stirring at room temperature. 0.1 M HCl (1.4 g, 0.14 mmol) was added to the surfactant solution under vigorous stirring at room temperature to partially neutralize the salt to produce free amino acid. After the mixture had been stirred for 1 h, a mixture of 1.40 g TEOS and 0.20 g TMAPS (50% in methanol) was added to the mixture with stirring at 22 °C. Then, the mixture was allowed to react at 22 °C under static conditions for 2 h. The chiral mesostructured product thus formed was cured at 80 °C for an additional 15 h. The products were recovered by centrifugal separation and dried at 60 °C. Both the anionic surfactant and the organics of the CSDA used were removed by calcination at 650 °C for 6 h.

Preparation of chiral metal wire

The chiral metal wires were prepared by two-step impregnation, oxidation/reduction procedures, using chiral mesoporous silica as a hard template^{11,12}. Typically, 0.15 g chiral mesoporous silica was impregnated with 1.8–2.6 mmol of metal precursors (Co(NO₃)₂·6H₂O and [Pt(NH₃)₄](NO₃)₂) in total. The Co and Pt precursor was activated while being slowly heated to 350 °C in a stream of O₂, and then reduced with stream of H₂ at 430 °C and 300 °C, respectively.

TEM image simulation

The ideal projected potential of the chiral material, which consists of uniform density walls and empty chiral pores, was constructed. TEM image simulation was performed by using the projected potential (kinematical scattering approximation) and incorporating an absorption effect. The applied objective lens aperture was so small (~0.5 nm⁻¹) that the simulated image should be regarded as a bright-field image.

Received 7 January; accepted 29 March 2004; doi:10.1038/nature02529.

1. Davis, M. E. New vistas in zeolite and molecular sieve catalysis. *Acc. Chem. Res.* **26**, 111–115 (1993).
2. Gier, T. E., Bu, X., Feng, P. & Stucky, G. D. Synthesis and organization of zeolite-like materials with three-dimensional helical pores. *Nature* **395**, 154–157 (1998).
3. Wang, Y., Yu, J., Guo, M. & Xu, R. {[Zn₂(HPO₄)₄][Co(dien)₂]}·H₂O: A zinc phosphate with multidirectional intersecting helical channels. *Angew. Chem. Int. Edn Engl.* **42**, 4089–4092 (2003).
4. Davis, M. E. Ordered porous materials for emerging applications. *Nature* **417**, 813–821 (2002).
5. Baiker, A. Chiral catalysis on solids. *Curr. Opin. Solid State Mater. Sci.* **3**, 86–93 (1998).
6. Che, S. *et al.* A novel anionic surfactant templating route for synthesizing mesoporous silica with unique structure. *Nature Mater.* **2**, 801–805 (2003).
7. Tracey, A. S. & Zhang, X. Investigation into the mechanism for generation of the helical axis in cholesteric lyotropic liquid crystals. *J. Phys. Chem.* **96**, 3889–3894 (1992).
8. Acharya, D. P. *et al.* Phase behavior and effect of enantiomerism on potassium N-dodecanoyl alaninate/water/decanol system. *J. Oleo. Sci.* **52**, 407–420 (2003).
9. Sakamoto, K. & Hatano, M. Formation of chiral aggregation of acylamino acid in solutions. *Bull. Chem. Soc. Jpn* **53**, 339–343 (1980).
10. Iwahashi, M. *et al.* Solvent effect on chiral aggregate formation of acylamino acids. *J. Oleo. Sci.* **51**, 705–713 (2003).
11. Shin, H. J., Ryoo, R., Liu, Z. & Terasaki, O. Template synthesis of asymmetrically mesostructured platinum networks. *J. Am. Chem. Soc.* **123**, 1246–1247 (2001).
12. Tian, B. *et al.* Facile synthesis and characterization of novel mesoporous and mesorelief oxides with gyroidal structures. *J. Am. Chem. Soc.* **126**, 865–875 (2004).

Supplementary Information accompanies this paper on www.nature.com/nature.

Acknowledgements This work was partially supported by Core Research for Evolutional Science and Technology (CREST) of JST Corporation, and by the Swedish Research Council (VR). S.C. thanks the Japan Society for the Promotion of Science for a postdoctoral fellowship to start this work at Yokohama National University.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.C. (chesa@sjtu.edu.cn).

Analysing the 1811–1812 New Madrid earthquakes with recent instrumentally recorded aftershocks

Karl Mueller¹, Susan E. Hough² & Roger Bilham^{1,3}

¹Department of Geological Sciences, University of Colorado, Boulder, Colorado 80309-0399, USA

²US Geological Survey, 525 South Wilson Ave, Pasadena, California 91030, USA

³CIRES University of Colorado, Boulder, Colorado 80309-0216, USA

Although dynamic stress changes associated with the passage of seismic waves are thought to trigger earthquakes at great distances, more than 60 per cent of all aftershocks appear to be triggered by static stress changes within two rupture lengths of a mainshock^{1–5}. The observed distribution of aftershocks may thus be used to infer details of mainshock rupture geometry⁶. Aftershocks following large mid-continental earthquakes, where background stressing rates are low, are known to persist for centuries^{7,8}, and models based on rate-and-state friction laws

provide theoretical support for this inference⁹. Most past studies of the New Madrid earthquake sequence have indeed assumed ongoing microseismicity to be a continuing aftershock sequence^{10–12}. Here we use instrumentally recorded aftershock locations and models of elastic stress change to develop a kinematically consistent rupture scenario for three of the four largest earthquakes of the 1811–1812 New Madrid sequence. Our results suggest that these three events occurred on two contiguous faults, producing lobes of increased stress near fault intersections and end points, in areas where present-day microearthquakes have been hitherto interpreted as evidence of primary mainshock rupture. We infer that the remaining New Madrid mainshock may have occurred more than 200 km north of this region in the Wabash Valley of southern Indiana and Illinois—an area that contains abundant modern microseismicity, and where substantial liquefaction was documented by historic accounts. Our results suggest that future large mid-plate earthquake sequences may extend over a much broader region than previously suspected.

Previous studies based on disparate lines of evidence have concluded that the first (NM1, 16 December 1811, magnitude $M = 7.3$) and last (NM3, 7 February 1812, $M = 7.5$) of the four principal New Madrid earthquakes respectively occurred on two well-defined faults, the northeast-striking Cottonwood Grove fault^{12,13} and the west-dipping Reelfoot thrust^{11,14–17} (Fig. 1). The smaller, second earthquake of the sequence (NM1-A, 16 December 1811, $M = 7.0$) may have occurred on either of these faults, or on a separate fault. The third earthquake (NM2, 23 January 1812, $M = 7.0$) has hitherto been interpreted as a strike-slip rupture on the northern limb of the New Madrid Seismic Zone (NMSZ).

To explore further a rupture scenario for the sequence, we calculated elastic stress changes produced by a sequence of fault ruptures to determine whether they are predicted to inhibit or enhance subsequent rupture of nearby target faults, with appropriate

sensing of displacement, surface area and orientation. We assumed that the initial shocks triggered subsequent ones and that triggering was caused entirely by static stress changes. In our models we assumed that the end points of the Reelfoot thrust are similar to the extent of fault-related folding of the ground surface in 1812 (ref. 15) that occurs above this fault for ~27 km northeast from its intersection with the Cottonwood Grove fault. Evidence for surface folding above the Reelfoot fault in 1812 includes eyewitness accounts of disruptions to the Mississippi River¹⁶, radio-carbon dating of deformed Holocene sediments¹⁸, and surface geomorphology^{17,19}. We assumed that the Cottonwood Grove fault extends southwest of its intersection with the Reelfoot fault¹ (Fig. 1).

Whereas there is little doubt about the locations of NM1 and NM3, the locations of NM1-A and NM2 remain conjectural. We used Coulomb theory⁵ to test the plausibility of published rupture scenarios¹ using recent estimates for magnitudes¹⁴ and fault geometry¹¹. In our initial model, we assumed 4 m of dextral slip on the Cottonwood Grove fault (NM1, Figs 1, 2a) and 1 m of dextral slip on a 25-km-long northern extension of the same fault (NM1-A). We assign NM2 to a 40-km-long dextral rupture on the northeast-trending arm of microseismicity at the northern end of the NMSZ (Fig. 1), as proposed by previous studies¹. The final event, NM3, is assumed to release 5 m of slip on the northwest-trending Reelfoot blind thrust fault^{18,19} (Fig. 1). We find this sequence scenario problematic because the first two events generate a lobe of reduced Coulomb stress for either dextral or thrust rupture in the region of NM2 (a decrease of 0.2–1.0 bar) that should inhibit faulting here (Fig. 2a).

We next examined alternative fault ruptures that might provide a kinematically consistent scenario for the sequence. Numerous model sequences were examined to test the influence of variations in rupture strike, slip, length, depth and sequence order on the prevailing stress field before each successive large shock (Fig. 2b, c). The results of these calculations suggest that the large thrust rupture on the Reelfoot fault (NM3) was probably triggered by a strike-slip rupture on the Cottonwood Grove fault (NM1), and either a shorter strike-slip rupture at its northern end (Fig. 2b) or a rupture of a short segment of the Reelfoot thrust (Fig. 2c). Our preferred scenario (Fig. 3) includes a 60-km Cottonwood Grove rupture (4 m of dextral slip, NM1), followed by a short rupture on the Reelfoot fault (1 m of reverse slip, NM1-A), and a large rupture on the Reelfoot fault (5 m of reverse slip, NM3).

No models satisfactorily include NM2 as a contiguous rupture, an issue we address in the following section. All consistent models reduce the geometry of the NMSZ to two main faults: the northeast-striking Cottonwood Grove and northwest-striking Reelfoot faults. Coulomb failure models derived from our preferred scenario include lobes of enhanced failure conditions (>1 bar) that correspond to limbs of microseismicity emanating from the three termination points of these two faults (Fig. 3). The lobe of microseismicity extending southeast from the intersection of the two faults has puzzled previous investigators because of its lack of surface expression¹⁹, and the requirement that should it represent a contiguous active fault, it must reverse its mechanism during the earthquake cycle²⁰.

The NM2 mainshock has been the most enigmatic of the four principal events because it was smaller than NM1 and NM3, and documented in fewer eyewitness accounts¹⁵. Using the method of Bakun and Wentworth²¹ we obtained an optimal location at -88.43° W, 36.95° N and a magnitude of 6.8 (Fig. 4). The location is not well constrained and can be almost equally well fitted by locations up to 100 km northwest or northeast of the optimal solution. Although the intensity distribution alone cannot rule out a source in the New Madrid region, the strongest shaking during this event was located significantly farther northeast than that from NM1-A, which was almost certainly in the New Madrid

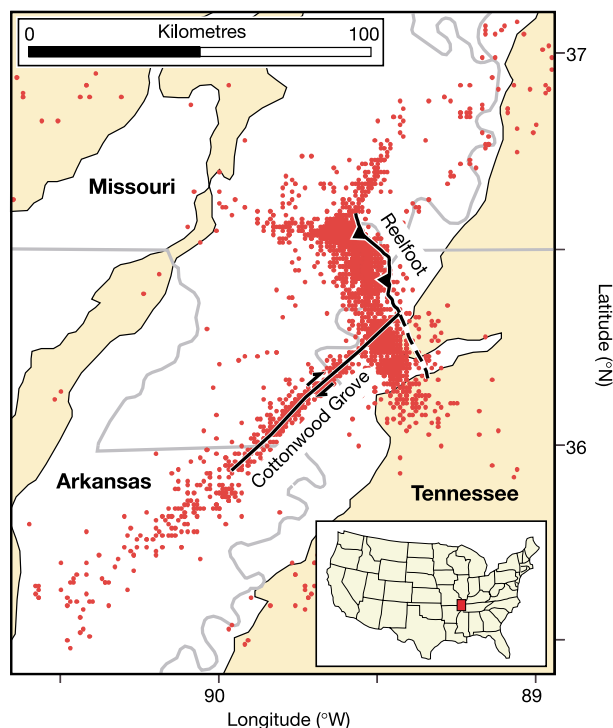


Figure 1 Map of the New Madrid seismic zone as illustrated by microseismicity between 1974 and 1996. The white region is the floodplain of the Mississippi River; grey lines are state boundaries. Locations of the Cottonwood Grove and Reelfoot faults are shown as solid black lines; the southeast extension of the Reelfoot thrust is shown as a dashed line.

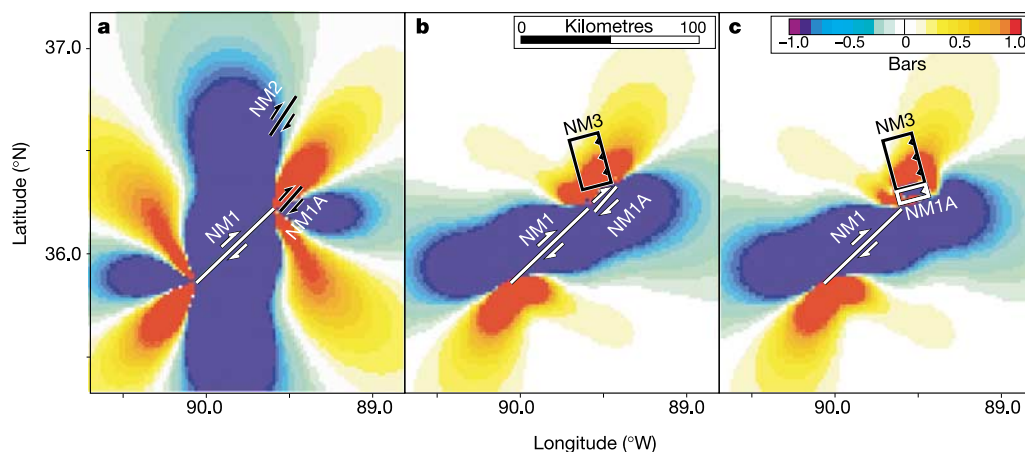


Figure 2 Sequential Coulomb failure stress associated with hypothesized ruptures. All models map stress changes at 12 km, with a coefficient of friction of 0.4 and a maximum principal stress oriented horizontally at N73°E. Earthquakes modelled shown as occurring on source faults (white lines). Black lines denote receiver faults that have not yet slipped during an earthquake. Slip on dextral faults is shown by arrows; barbs denote thrusts. **a**, Coulomb failure enhancement resulting from 4 m of N46°E dextral-shear following rupture during NM1 of the Cottonwood Grove fault for vertical strike-slip faults oriented N34°E. This brings a subsequent event at the northeast end of the

Cottonwood Grove fault that might fail during NM1-A closer to failure, but retards failure near the traditional location of NM2. Failure of a dextral fault during NM2 is also inhibited if NM1-A is modelled as a short segment of the Reelfoot thrust. **b**, The Reelfoot thrust (NM3) shown as being brought closer to failure by 4 m of dextral slip during NM1 and 1 m of slip in NM1-A. **c**, The Reelfoot thrust is also brought closer to failure by 4 m of dextral slip during NM1 followed by 1 m of reverse slip on the Reelfoot thrust from NM1-A.

region^{1,22}. The intensity distribution suggests that NM2 was a remotely triggered earthquake.

The absence of a high-intensity ‘bulls-eye’ for NM2 (Fig. 4) is consistent with a location in southeastern Illinois, a region that was very sparsely populated in 1812 (ref. 23). Moreover, had the 1812 event resembled the 1968 $M_b = 5.3$ southern Illinois earthquake, which occurred at 20–25 km depth²⁴, it would have produced a broad intensity distribution. A location north and east of NM1 and NM3 would explain why remotely triggered earthquakes occurred in northern Kentucky after NM2 (ref. 25), but none were observed following NM1. A southeastern Illinois location could also explain reported liquefaction accounts there²³ (Fig. 4). Our rupture scenario for NM1, NM1-A, and NM3 is also consistent with the observation of three liquefaction events in the New Madrid region in 1811–1812 (ref. 26).

One historic account provides an intriguing suggestion of a possible surface rupture some 220 km from New Madrid. This account, by Yearby Land, described “a big crack [that] was made in the ground” with two feet of vertical displacement to the south²³. Even in 1858, the feature (38.07°N, 88.11°W) could be traced for a reported distance of two miles. Near this crack Land stated that “piles and piles of pure, snow white sand were heaved up” including some as big as “several wagon loads”. Field reconnaissance by the authors verified many of the details in the Land account, and confirmed evidence of sand blows on the surface of the field where they were reported. In addition to providing a clear account of liquefaction, this report appears to describe either surface rupture on an east–west trending fault or substantial ground slumping. None of the earthquakes caused significant damage in the region beyond toppling chimneys, but, in addition to the sparse population, it has been noted that the buildings were all extremely flexible²³. Land’s account does not link the effects to a particular earthquake, and thus has not been used in previous intensity studies. At a minimum, however, it indicates that ground motion severe enough to cause substantial liquefaction occurred during the sequence, in a location centred in the area of highest shaking from NM2. We also note that a linear northwest-trending belt of rapidly occurring microseismicity coincides with this location²⁷, and that

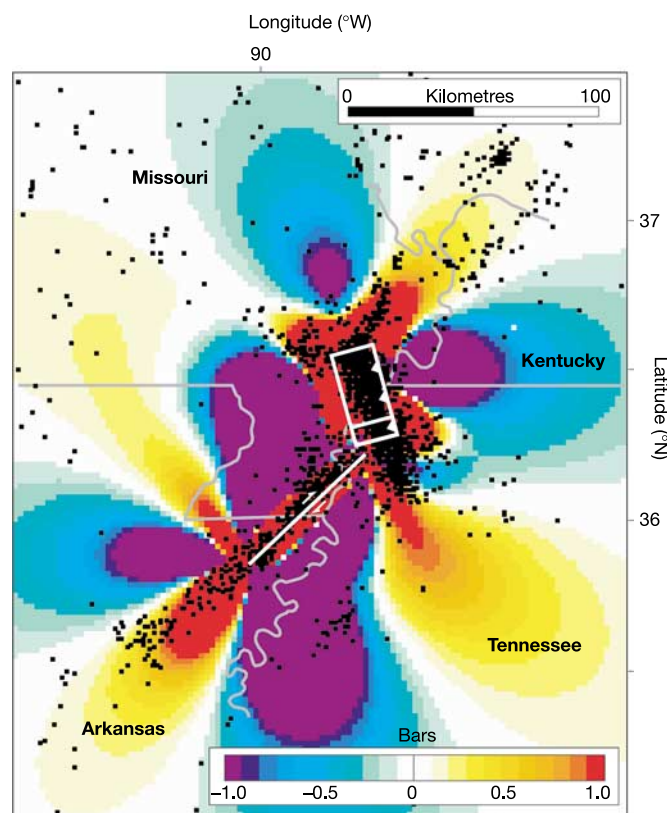
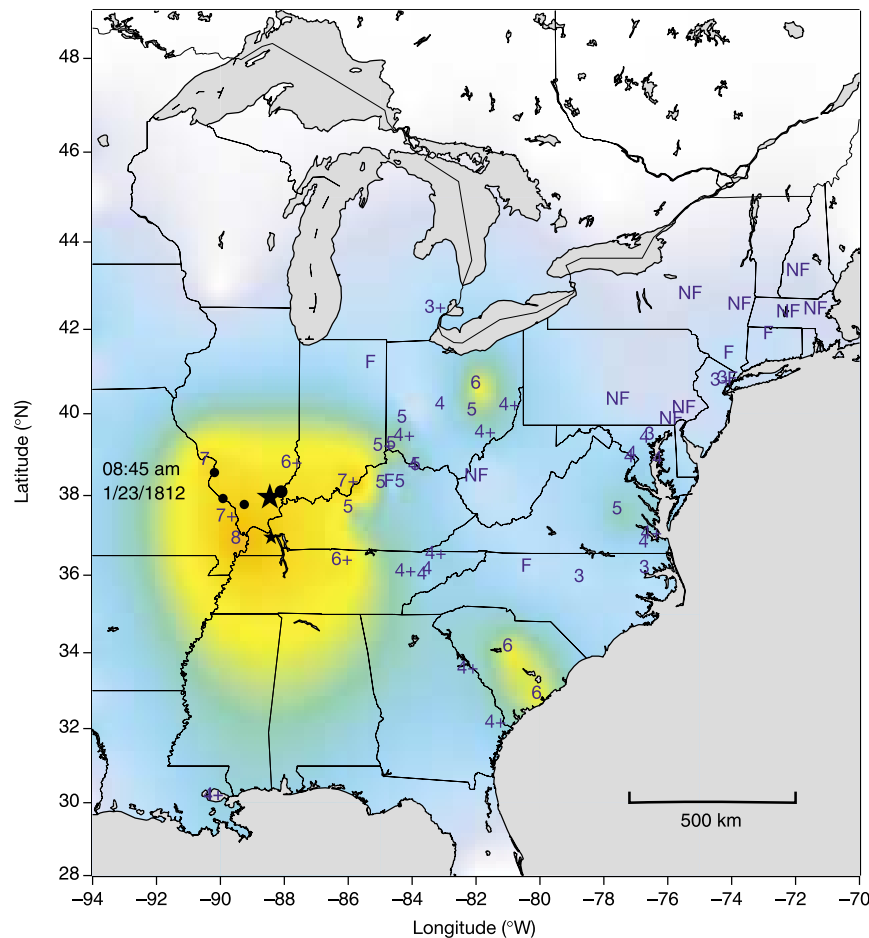


Figure 3 Elastic stress field at 12 km depth produced by NM1, NM1A and NM3 in 1811–1812. Preferred fault ruptures are shown as white lines (NM1, 4 m dextral slip on the 60-km-long Cottonwood Grove fault); a small white rectangle (NM1-A, 1 m reverse slip on Reelfoot thrust); and a large rectangle (NM3, 5 m reverse slip rupture of the Reelfoot thrust). Note the general coincidence of arms of microseismicity with lobes of enhanced stress.



Perceived shaking	Not felt	Weak	Light	Moderate	Strong	Very strong	Severe	Violent	Extreme
Potential damage	None	None	None	Very light	Light	Moderate	Moderate/heavy	Heavy	Very heavy
Peak acceleration (% of g)	<0.17	0.17–1.4	1.4–3.9	3.9–9.2	9.2–18	18–34	34–65	65–124	>124
Peak velocity (cm s^{-1})	<0.1	0.1–1.1	1.1–3.4	3.4–8.1	8.1–16	16–31	31–60	60–116	>116
Instrumental intensity	I	II–III	IV	V	VI	VII	VIII	IX	X+

Figure 4 Mercalli intensities inferred for NM2, 23 January 1812, with symbols centred on towns. To map the distribution of shaking we use simple mathematical interpolation¹⁰. The small star shows the optimal epicentre using quantitative regression²¹; the large star shows the location of the 1968 Southern Illinois earthquake. Black circles correspond to

locations where liquefaction was documented during the 1811–1812 sequence. The easternmost and largest circle identifies the location of the White County, Illinois, event, the account of which describes substantial liquefaction as well as surface cracking²³.

the rate of $M > 3$ earthquakes has previously been quite high in the Wabash Valley.

With the assumption that the principal earthquakes of the 1811–1812 New Madrid sequence constitute a progressively triggered sequence, we conclude that primary faulting in the NMSZ occurred as three ruptures on a linked strike-slip and thrust fault system. The third mainshock, hitherto assigned to a lobe of microseismicity at the north end of the thrust fault, may have been a triggered earthquake more than 200 km northeast of New Madrid in a region of abundant ongoing microseismicity. We interpret diffuse arms of microseismicity now occurring at the end points and intersections of the two main New Madrid faults as aftershocks occurring in regions where stress was increased by the mainshocks. Our interpretation is consistent with established magnitude/fault area results, and do not require exceptionally large fault areas or stress drop values for the New Madrid mainshocks. Ironically, however, the results of this study imply that the hazard associated with mid-continent earthquakes might be more widely distributed than

previously recognized. That is, large dynamically triggered remote earthquakes may follow large mid-plate mainshocks. □

Received 24 October 2003; accepted 7 April 2004; doi:10.1038/nature02557.

- Johnston, A. C. & Schweig, E. S. The enigma of the New Madrid earthquakes of 1811–1812. *Annu. Rev. Earth Planet. Sci. Lett.* **24**, 339–384 (1996).
- Kilb, D., Gomberg, J. & Bodin, P. Triggering of earthquake aftershocks by dynamic stresses. *Nature* **408**, 570–574 (2000).
- Das, S. & Scholz, C. H. Off-fault aftershock clusters caused by shear-stress increase. *Bull. Seismol. Soc. Am.* **71**, 1669–1675 (1981).
- Reasenber, P. A. & Jones, L. M. Earthquake hazard after a mainshock in California. *Science* **243**, 1173–1176 (1989).
- Toda, S., Stein, R. S., Reasenber, P. A., Dieterich, J. H. & Yoshida, A. Stress transferred from the 1995 Mw = 6.9 Kobe Japan, shock: Effect on aftershocks and future earthquake probabilities. *J. Geophys. Res.* **103**, 24,543–24,565 (1998).
- Seeber, L. & Armbruster, J. G. Earthquakes as beacons of stress change. *Nature* **407**, 69–72 (2000).
- Ebel, J., Bonjer, E. & Oncescu, M. Paleoseismicity: Seismicity evidence for past large earthquakes. *Seismol. Res. Lett.* **71**, 283–294 (2000).
- Gan, W. & Prescott, W. H. Crustal deformation rates in central and eastern U.S. inferred from GPS. *Geophys. Res. Lett.* **28**, 3733–3736 (2001).
- Dieterich, J. A constitutive law for rate of earthquake production and its application to earthquake clustering. *J. Geophys. Res.* **99**, 2601–2618 (1994).

10. Nuttli, O. W. The Mississippi Valley earthquakes of 1811 and 1812: Intensities, ground motion, and magnitudes. *Bull. Seismol. Soc. Am.* **63**, 227–248 (1973).
11. Mueller, K. & Pujol, J. Three-dimensional geometry of the Reelfoot blind thrust: implications for moment release and earthquake magnitude in the New Madrid seismic zone. *Bull. Seismol. Soc. Am.* **91**, 1563–1573 (2001).
12. Johnston, A. C. Seismic moment assessment of earthquakes in stable continental regions III, New Madrid 1811–1812, Charleston 1886, and Lisbon 1755. *Geophys. J. Int.* **126**, 314–344 (1996).
13. Guccione, M. J., VanArsdale, R. B. & Hehr, L. H. Origin and age of the Manila high and associated Big Lake “sunklands” in the New Madrid seismic zone, northeastern Arkansas. *Geol. Soc. Am. Bull.* **112**, 579–590 (2000).
14. Hough, S. E., Armbruster, J. G., Seeber, L. & Hough, J. F. On the modified Mercalli intensities and magnitudes of the 1811–1812 New Madrid, central United States earthquakes. *J. Geophys. Res.* **105**, 23839–23864 (2000).
15. Baldwin, J. N., Barron, A. D., Kelson, K. I., Harris, J. B. & Cashman, S. Preliminary paleoseismic and geophysical investigation of the north Farrenburg lineament: Primary tectonic deformation associated with the New Madrid North fault? *Seismol. Res. Lett.* **73**, 393–413 (2002).
16. Odum, J. K., Stephenson, W. J., Shedlock, K. M. & Pratt, T. L. Near-surface structural model for deformation associated with the February 7, 1812, New Madrid, Missouri, earthquake. *Geol. Soc. Am. Bull.* **110**, 149–162 (1998).
17. Van Arsdale, R. B., Kelson, K. I. & Lumsden, C. H. Northern extension of the Tennessee Reelfoot scarp into Kentucky and Missouri. *Seismol. Res. Lett.* **65**, 57–62 (1995).
18. Kelson, K. I., Simpson, G. D., VanArsdale, R. B., Haraden, C. C. & Lettis, W. R. Multiple late Holocene earthquakes along the Reelfoot fault, central New Madrid seismic zone. *J. Geophys. Res.* **101**, 6151–6170 (1992).
19. Guccione, M. J. *et al.* Stream response to repeated coseismic folding, Tiptonville Dome, western Tennessee. *Geomorph.* **43**, 313–349 (2002).
20. Gombert, J. & Ellis, M. Topography and tectonics of the central New Madrid seismic zone: Results of numerical experiments using a three-dimensional boundary element program. *J. Geophys. Res.* **99**, 20299–20310 (1994).
21. Bakun, W. H. & Wentworth, C. M. Estimating earthquake location and magnitude from seismic intensity data. *Bull. Seismol. Soc. Am.* **87**, 1502–1521 (1997).
22. Hough, S. E. & Martin, S. Magnitude estimates of two large aftershocks of the 16 December 1811 New Madrid earthquake. *Bull. Seismol. Soc. Am.* **92**, 3259–3268 (2002).
23. Berry, D. L. The Illinois earthquake of 1811 and 1812. *Illinois State Hist. Soc. Trans.* **12**, 74–78 (1908).
24. Stauder, W. & Nuttli, O. W. Seismic studies—south-central Illinois earthquake of November 9, 1968. *Bull. Seismol. Soc. Am.* **60**, 973–981 (1970).
25. Hough, S. E. Triggered earthquakes and the 1811–1812 New Madrid, central U.S. earthquake sequence. *Bull. Seismol. Soc. Am.* **91**, 1574–1581 (2001).
26. Tuttle, M. P. The use of liquefaction features in paleoseismology: Lessons learned in the New Madrid seismic zone, central United States. *J. Seismol.* **5**, 361–380 (2001).
27. Pavlis, G. L., Rudman, A. J., Pope, B. P., Hamburger, M. W. & Bear, G. W. Seismicity of the Wabash Valley seismic zone based on a temporary seismic array experiment. *Seismol. Res. Lett.* **73**, 751–761 (2002).

Acknowledgements Support for this work was provided to R.B. by the National Science Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to K.M. (Karl.Mueller@colorado.edu).

Articulated Palaeozoic fossil with 17 plates greatly expands disparity of early chitons

Michael J. Vendrasco¹, Troy E. Wood³ & Bruce N. Runnegar^{1,2*}

¹Department of Earth and Space Sciences, ²Institute of Geophysics and Planetary Physics and Molecular Biology Institute, University of California, Los Angeles, California 90095, USA

³Department of Biology, JH 142, 1001 E. Third Street, Indiana University, Bloomington, Indiana 47405, USA

* Present address: NASA Astrobiology Institute, MS 240-1, Ames Research Center, Moffett Field, California 94035, USA

Modern chitons (Mollusca: Polyplacophora) possess a highly conserved skeleton of eight shell plates (valves) surrounded by spicules or scales, and fossil evidence suggests that the chiton skeleton has changed little since the first appearance of the class in the Late Cambrian period (about 500 million years before present, Myr BP). However, the Palaeozoic problematic taxon

Multiplicaphora^{1–5}, in spite of having a more complex skeleton, shares several derived characters with chitons. The enigmatic status of the Multiplicaphora is due in part to the fact that its members had an exoskeleton of numerous calcium carbonate valves that usually separated after death. A new articulated specimen from the Carboniferous period (about 335 Myr BP) of Indiana reveals that multiplicaphorans had a dorsal protective surface composed of head and tail valves, left and right columns of overlapping valves (five on each side), and a central zone of five smaller valves, all surrounded by an annulus of large spines. Here we describe and name the articulated specimen and present evidence that multiplicaphorans were chitons. Thus the highly conserved body plan of living chitons belies the broad disparity of this clade during the Palaeozoic era.

The modern chiton skeleton consists of one column of eight calcareous shell plates (valves) surrounded by a ring of scales or spicules. Articulated fossil chitons share this skeletal body plan. Moreover, isolated fossil valves assigned to the Polyplacophora show bilateral symmetry, consistent with a single column of valves. Consequently, the skeletal morphology of this clade has been interpreted as evolutionarily static, with little change since its origin in the Cambrian⁶.

A new, exceptionally preserved multiplicaphoran (Fig. 1a) from the Carboniferous of Indiana allows us to link this group with the chitons, and reveals that this clade exhibited far greater skeletal variation during the Palaeozoic than is generally believed. Multiplicaphorans (Devonian to Permian, ~380–265 Myr BP, fossils from North America) have been considered problematic in part because the detailed nature of their skeleton was unknown. The new fossil reveals that the multiplicaphoran skeleton consists of three longitudinal columns of shell plates plus a head and tail plate, surrounded by elongate, hollow spines (Fig. 1c), an arrangement similar to that of a recently described Ordovician chiton *Echinochiton dufoei*⁷. The *Echinochiton* skeleton consists of a central column of eight valves, two flanking columns of small, dorsally projecting ‘scutes’, and a girdle harbouring eight pairs of elongate, hollow spines that grew by marginal accretion. Thus, the skeleton of *Echinochiton* is intermediate in form between those of multiplicaphorans and typical chitons.

Multiplicaphorans share with other chitons an iterated sequence of overlapping, calcareous, marginally accreted valves, surrounded by an armoured girdle. In both groups the anterior and posterior valves are morphologically distinct from the intervening valves; the valve surfaces are porous, and an inner shell layer (articulamentum) projects from the growing margins of valves (Fig. 2a, c). In addition, microscopic canal systems occur within the valves of both taxa (Fig. 2b, d). These tubular canals exist at the interface of two shell layers; numerous, equally spaced canals branch off obliquely towards the surface, and, in certain areas, vertical canals pass through all shell layers^{5,8}. In chitons, these canals house tissues that connect sensory structures, known as aesthetes, at the valve surface to the body below⁸. We envisage a similar function for the canals in multiplicaphoran valves.

Despite these similarities, multiplicaphorans are distinct from typical chitons. Multiplicaphorans have three anterior–posterior columns of valves and a sevenfold rather than eightfold iteration of major skeletal elements. In addition, the large, hollow, lateral spines of multiplicaphorans are probably not homologous with the small, solid, girdle spicules of chitons, and, conversely, they share many similarities with chiton valves^{5,9}. These similarities, which include a porous surface, internal canals, projections of the articulamentum and growth by marginal accretion, suggest that the spines may instead be homologous with chiton valves.

An elementary cladistic analysis strongly supports the placement of multiplicaphorans within the total group Polyplacophora (Fig. 3). All modern chiton valves possess projections of the articulamentum (articulaments) that serve to anchor the valves in

the fleshy girdle. Early Palaeozoic chitons such as *Matthevia* and *Chelodes* lack such projections¹⁰, and so are considered to be members of the stem group. Articulaments evolved before the last common ancestor of all living chitons, so they do not necessarily identify a member of the chiton crown group, but they do distinguish relatively derived chitons from primitive ones. We therefore use a complete Carboniferous chiton with well-developed articulaments (*Glaphurochiton*) to place a lower limit on the depth of the chiton crown group (Fig. 3). *Glaphurochiton* and its late Palaeozoic relatives are probably derived members of the stem group, because some exhibit features that are not found in any living chiton.

Echinochiton also lacks any indication of articulaments, and as its valves are similar to those of *Matthevia*⁷, it falls within the chiton stem group. Multiplacophorans are more problematical, because they possess articulaments but lack the conserved eight-valve array of typical chitons. Nevertheless, they also lie firmly within the polyplacophoran stem group in this cladistic analysis (Fig. 3). Whether or not *Echinochiton* and the Multiplacophora are monophyletic remains unresolved because the bootstrap support for the node leading to *Polysacos* gen. nov. and *Echinochiton* (Fig. 3) is weak (49%).

The Silurian vermiform mollusc *Acaenoplax hayae*^{11,12} had an aplacophoran-shaped body, serially arranged bands of calcareous spines and seven (of eight?) chiton-like valves. *Acaenoplax* almost certainly belongs to the aplacophoran stem group¹² (Fig. 3), although it is possible that Aplacophora is a grade¹³ rather than a clade¹⁴. *Halkieria evangelista*^{15,16}, a slug-shaped Cambrian animal covered with thousands of hollow mineralized sclerites, plus head and tail valves, was also included in the analysis. Despite similarities to chitons, *H. evangelista* was assigned to the brachiopod stem group when first described¹⁶ because some living brachiopods pass through a halkieriid-like form early in their development¹⁷. Also, *Micrina* and *Tannuolina*, whose isolated valves have been interpreted as being derived from halkieriids, have valve microstructure similar to that of early brachiopods¹⁸. For these reasons, we used the

living brachiopods *Lingula* and *Novocrania* as an outgroup. In our preferred tree (Fig. 3), *Halkieria* groups with the Mollusca rather than the Brachiopoda. The same result is recovered even when its (unknown) valve composition is coded as phosphatic, the preferred character state for the stem-group brachiopod hypothesis. Although the exact placement of *Halkieria* is unclear, it seems likely that the valves of all the disparate taxa analysed here, irrespective of differences in shell microstructure, are fundamentally homologous.

The cladistic analysis suggests that aplacophorans and chitons are sister groups within a monophyletic Aculifera (Aplacophora plus Polyplacophora). In contrast, a recent cladistic analysis of molluscs, incorporating a wider range of morphological characters but no fossils, portrays chitons as gradational between aplacophorans and conchiferans (shelled molluscs)¹³. However, the position of chitons within the Mollusca may not bear on the origin of multiplacophorans. The greater complexity of the multiplacophoran skeleton compared with that of typical chitons is probably a derived feature of the group. The presence of small scutes as well as large central valves in *Echinochiton* suggests that the multiplacophoran exoskeleton may have evolved from a chiton-like progenitor through the insertion of new left and right sets of valves and the concomitant reduction of the central valves. This hypothesis is supported by the observation that multiplacophoran valves have articulaments, which are absent in pre-Devonian chitons. Also, the oldest known multiplacophorans⁵ are from the Devonian, compared with the Cambrian appearance of the earliest known members of the stem group¹⁰.

The principal alternative possibility, that the multiplacophoran skeleton was primitive, would imply that the sevenfold iteration of major skeletal elements is plesiomorphic for chitons. This hypothesis is supported by observations of a sevenfold iteration of bare regions between bundles of dorsal carbonate spicules in a postlarval aplacophoran¹⁴, the seven-valved *Acaenoplax* (but see ref. 12), and a seven-valved state seen in the early ontogeny of chitons¹⁹. Thus, the sevenfold iteration may represent the primitive

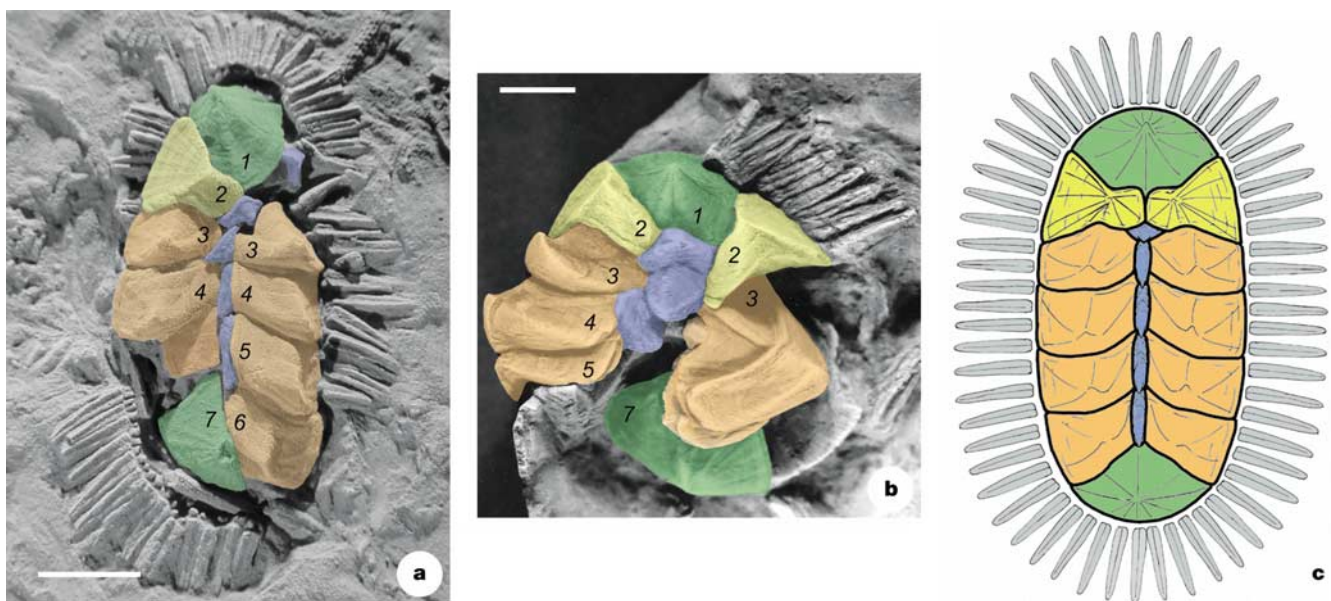


Figure 1 Articulated multiplacophorans and a reconstruction of the skeleton. **a**, Holotype of *Polysacos vickersianum* gen. et sp. nov. **b**, Holotype of *Strobilepis spinigera* Clarke 1888. **c**, Reconstruction of the complete dorsal exoskeleton of a multiplacophoran, colour-coded as in **a** and **b**. Head valves are green, anterior pair of intermediate lateral

valves are yellow, other intermediate valves are orange and central valves are blue. In **a**, the unnumbered left lateral valve that is partially obscured by valve 4 is valve 6; valve 5 is missing. In **b**, a small portion of another lateral valve is visible between valves numbered 3 and 7. Scale bars, 5 mm.

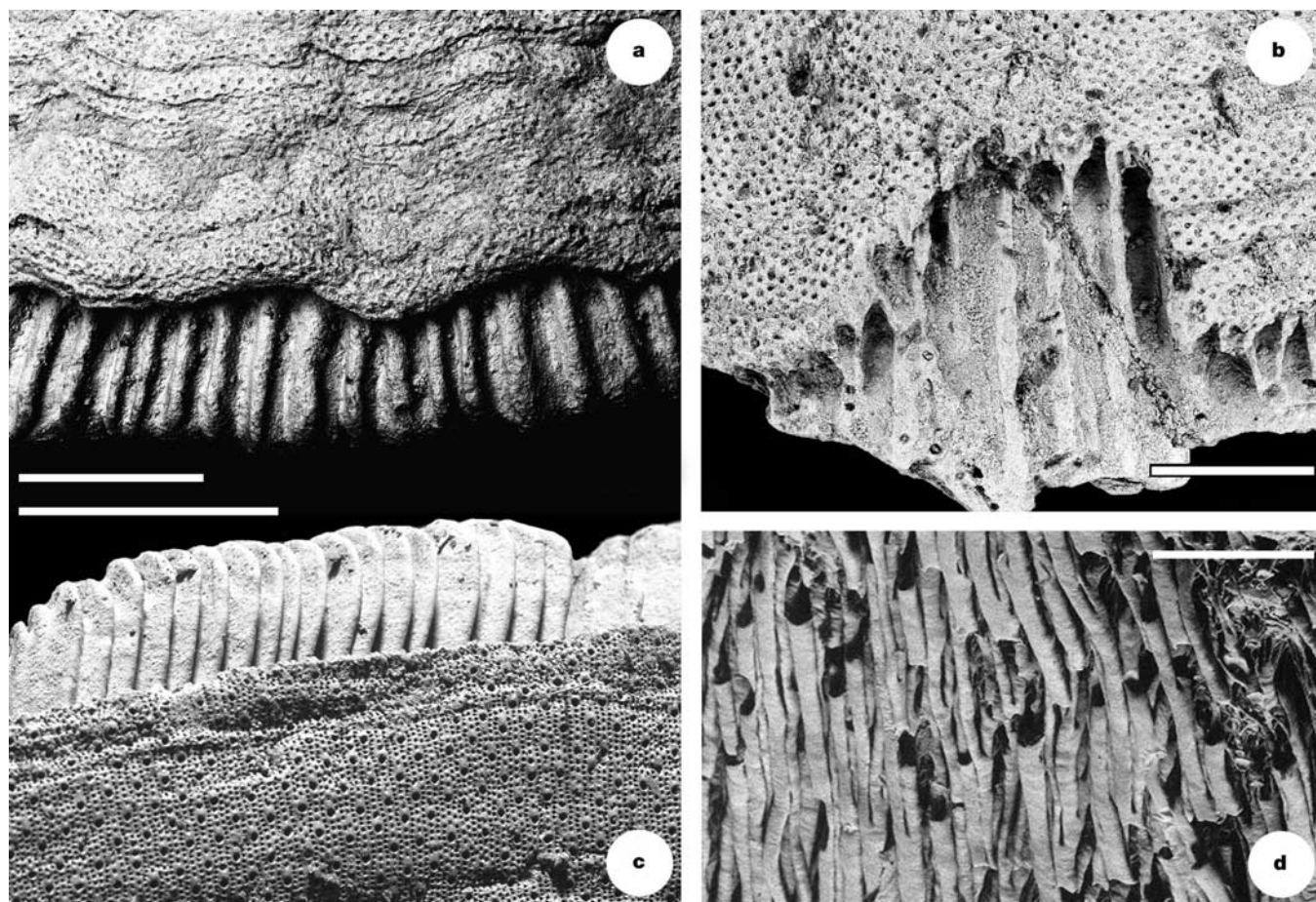


Figure 2 Fine structure of multiplacophoran and chiton valves. **a**, Edge of intermediate valve of *Diadeloplax paragrapsima*, Pennsylvanian, Oklahoma, showing porous valve surface, comarginal growth increments and the projecting inner layer (articulamentum) that was embedded in the body wall. **b**, Broken edge of intermediate valve of *D. paragrapsima*, Pennsylvanian, Oklahoma, showing porous surface of outer layer and possible casts of aesthete canals in lower layer. **c**, Edge of

intermediate valve of living *Chiton tulipa* showing porous valve surface, comarginal growth increments and projecting inner layer (articulamentum) that was embedded in the body wall. **d**, Epoxy casts of aesthete canals in intermediate valve of the living chiton, *Mopalia muscosa*, for comparison with **b**. Scale bars: 500 μm (**c**); 250 μm (**b**); 200 μm (**d**); 1 mm (**a**).

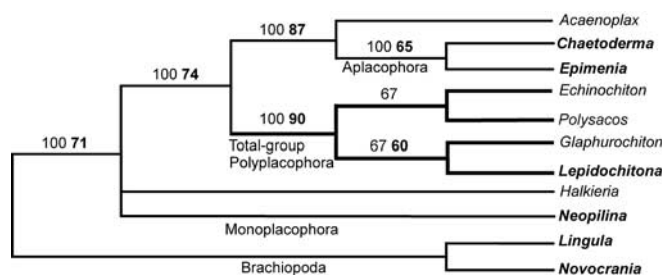


Figure 3 Cladogram showing placement of the Multiplacophora (Polysacos) and *Echinochiton* within the Polyplacophora. The analysis was based on 14 characters (see Supplementary Table 1 for details). The tree shown is the 50% majority rule consensus of the nine shortest trees (of length 24) obtained from an exhaustive search using maximum parsimony. Numbers not in bold are the percentages of the nine trees that support the nodes, whereas the numbers in bold are percentages of bootstrap support based on 10,000 replicates. Bootstrap values lower than 50% were not included in the diagram. Taxon names in bold type refer to living animals, all others are extinct. The analysis was carried out with PAUP 4.0 (ref. 25).

calcifying condition of the Aculifera or, perhaps, of the Mollusca in general.

Although it is difficult to polarize the transition between typical chitons and multiplacophorans, the homologies established here demonstrate that multiplacophorans were chitons. Thus, polyplacophorans exhibited greater morphological disparity during the Palaeozoic than they do today. Multiplacophorans, like trilobites, were pruned from the metazoan tree before the radiation of the modern fauna during the Mesozoic. Subsequent diversification within the Polyplacophora yielded only modest variation on one general theme. □

Methods

Systematic palaeontology

Phylum Mollusca Cuvier 1797
Class Polyplacophora de Blainville 1816
Order Multiplacophora Hoare and Mapes 1995

Diagnosis. The dorsal exoskeleton consists of a head and tail valve and three longitudinal columns of intermediate valves (five per column), all surrounded by a girdle harbouring outwardly projecting, hollow, calcareous spines. The left and right columns of valves have mirror symmetry; valves of the central column are smaller and bilaterally symmetrical. There are 17 valves, all of which are thick, calcareous, overlap and grew by marginal accretion. Both valves and spines have pores at the surface that extend downwards into internal (aesthete) canals. There are two main shell layers. The outer layer (tegumentum) of each valve has numerous canals whereas the inner layer (articulamentum) has fewer canals and projects beyond the outer layer at the growing margins. In the outer layer, the canals

course vertically from the dorsal valve surface to the boundary with the inner layer. At the boundary larger canals course horizontally towards the valve margins. In a few places the canals penetrate both shell layers.

Discussion. The class Multiplacophora was originally based on a partly articulated fossil⁵ (Fig. 1b), named *Strobilepis* by Clarke in 1888, as well as a number of disarticulated shell plates and spines^{5,9} placed in the genus *Diadeloplax*. A reconstruction of *Strobilepis* showing three anterior–posterior columns of shell plates was provided with the original description of *Strobilepis spinigera*¹, but this reconstruction was called into question recently and an alternative reconstruction having a single line of 12 or more plates was proposed^{5,9}. The new specimen, *Polysacos vickersianum* gen. et sp. nov., confirms Clarke's original reconstruction, demonstrates that multiplacophorans had a total of 17 dorsal valves, and preserves the previously unrecognized tail valve (Fig. 1a, c). *Strobilepis* Clarke was originally described as a barnacle^{1,2}.

Two similar species, *Protobalanus hamiltonensis*²⁰ and *Hercolepas signata*²¹, known from articulated but poorly preserved specimens, were also originally considered to be barnacles. These taxa were subsequently placed in their own order, the Hercolepadida⁴, within the class Machaeridia, an extinct group of uncertain affinities. This assessment was later questioned and the hercolepadids were removed from the Machaeridia²². Similarities in the nature of the valves and spines between hercolepadids and multiplacophorans have been noted, but these two groups have not been formally united^{5,23}. The entire skeleton of *H. signata* is not known, but it clearly had multiple columns of shell plates surrounded by a ring of projecting spines, each with a shallow, longitudinal groove on the dorsal surface^{2,21}. A similar groove is seen in the spines of multiplacophorans. The valve surface of *Hercolepas* is ornamented with punctae², another similarity with multiplacophorans. The two known specimens of *P. hamiltonensis* suggest that this species had a comparable arrangement of 17 valves encircled by spines²⁴. Because of these similarities in skeletal anatomy, the hercolepadids and multiplacophorans are here grouped as members of the Multiplacophora. We retain the younger name because it is more appropriate as a higher taxon of the Mollusca.

Genera included in the Multiplacophora are *Strobilepis*, *Diadeloplax*, *Aenigmactectus*, *Protobalanus*, *Hercolepas* and *Polysacos* gen. nov. The stratigraphic range is from Devonian to Permian.

Polysacos gen. nov. Vondrasco, Wood and Runnegar

Etymology. From *poly* (Greek, meaning many) and *sakos* (Greek, meaning shield), neuter. **Type species.** *Polysacos vickersianum* sp. nov.

Diagnosis. As for order, but with valves ornamented with low radial ridges and grooves. Marginal spines have a concave upper surface, a porous external layer and two peg-like projections of the underlying shell layer on the proximal end. Valves are ornamented with radiating ridges and faint growth lines. Head valve is suboval in dorsal view, arched, slightly longer than wide and with distinct mucro. The anterior-most lateral valve has a flat, triangular lateral surface and is as long as it is wide. The remaining lateral valves are suboval in dorsal view, highly arched, slightly wider than long and with a distinct apex. The central valves range from squarish in dorsal view (near the head valve) to spinose (near the tail valve). The tail valve has a distinct mucro close to the rounded posterior margin.

Discussion. *Polysacos* differs from *Strobilepis*^{1,5} in that the apices of lateral intermediate valves are much lower, are not markedly curved and are located closer to the posterior end of the valve. The central valves are narrower in *Polysacos* than in *Strobilepis*. *Polysacos* differs from *Diadeloplax*^{5,9} in being less coarsely ornamented, in having peg-like projections of the lower shell layer in the spines, and in having broad comarginal ridges along the posterior margins of the lateral intermediate valves. *Polysacos* differs from *Protobalanus*^{19,24} and *Hercolepas*²¹ in its considerably larger size, more elongate body and in having longer lateral spines. *Polysacos* differs from *Aenigmactectus*²³ in having much shorter and simpler articulation projections in valves and spines.

Polysacos vickersianum sp. nov. Vondrasco, Wood and Runnegar

Etymology. Named for Donna Vickers.

Holotype. USNM 520742, Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Washington DC, USA.

Locality. Edwardsville Formation, Carboniferous (early Mississippian), Monroe County, Indiana, USA.

Diagnosis. As for genus. The holotype is an incompletely articulated individual, 25 mm in total length, preserved in near-life orientation. In addition to head and tail valves, seven of the ten lateral intermediate valves are well preserved and the middle valve of each series is present on both sides (valve 4, Fig. 1a). The anterior lateral valve is missing from the right side, the fifth lateral valve is missing from the left side, and the sixth lateral valve of the left side lies beneath valve 4 on the left side. Four of the five intermediate valves on each side are similar in shape; the fifth, which is the anterior-most one, is different and, remarkably, it overlaps both the head valve as well as the intermediate valve that follows it. Every other intermediate valve overlaps the one behind it. The apices of the head and tail valves are at the anterior and posterior ends of their respective valves, whereas the apices of the intermediate valves are centrally positioned and curve slightly forward. Four of the five central valves are in place along the dorsal midline but the anterior one is displaced so that it lies near the right margin of the head valve. Many marginal spines are missing, but it is clear from those that are present that there were approximately 60 spines in all. The spines have concave dorsal surfaces and prominent projections at their proximal ends that presumably served to anchor them in the body wall.

Discussion. When Hoare and Mapes² re-described the holotype of *S. spinigera* they noted that the valves labelled 3–5 in Fig. 1b appeared to have been reversed with respect to the body after death. They came to this conclusion because the order of valve overlap, coupled with the direction of curvature of the valve apices, indicated that the spine-bearing end of the specimen was anterior whereas the disposition of the other valves suggested the opposite. However, the holotype of *P. vickersianum* has intermediate and

terminal valves arranged in exactly the same way (Fig. 1a), so it is clear that this was the way the valves were arranged in both animals during life. We consider the valves labelled 1 in Fig. 1a, b to be the anterior (head) valves because the majority of valves are imbricated in that direction. Only the first intermediate valve (2, Fig. 1a, b) is extraordinary in that it overlaps both preceding and succeeding valves. This anomalous situation, plus the terminal position of the apex of the head valve, is not seen in other chitons.

Computerized tomography imaging

The specimen was imaged in three dimensions using the X-ray computerized tomography facility at the University of Texas, Austin (Supplementary Fig. 1).

Received 1 December 2003; accepted 5 April 2004; doi:10.1038/nature02548.

- Hall, J. & Clarke, J. M. Trilobites and other Crustacea of the Oriskany, Upper Helderberg, Hamilton, Portage, Chemung and Catskill Groups. *Geol. Surv. NY Palaeontol.* **7**, 1–236 (1888).
- Withers, T. H. Some Palaeozoic fossils referred to the Cirripedia. *Geol. Mag.* **6**, 112–123 (1915).
- Hoare, R. D. *Lobarochiton*, new generic designation for *Gryphochiton*? *Anomolous* (Rowley) (Polyplacophora). *Ohio J. Sci.* **76**, 116–118 (1976).
- Dzik, J. in *Problematic Fossil Taxa* (eds Hoffman, A. & Nitecki, M. H.) 116–134 (Oxford Univ. Press, New York, 1986).
- Hoare, R. D. & Mapes, R. H. Relationships of the Devonian *Strobilepis* and related Pennsylvanian problematica. *Acta Palaeontol. Pol.* **40**, 111–128 (1995).
- Eernisse, D. J. & Reynolds, P. D. in *Microscopic Anatomy of Invertebrates* Vol. 5 *Mollusca I* (eds Harris, F. W. & Kohn, A. J.) 55–110 (Wiley Liss, New York, 1994).
- Pojeta, J. Jr, Eernisse, D. J., Hoare, R. D. & Henderson, M. D. *Echitonichiton dufouei*: a new spiny Ordovician chiton. *J. Paleontol.* **77**, 646–654 (2003).
- Baxter, J. M. & Jones, A. M. Valve structure and growth in the chiton *Lepidochitona cinereus* (Polyplacophora: Ischnochitonidae). *J. Mar. Biol. Assoc. UK* **61**, 65–78 (1981).
- Hoare, R. D. & Mapes, R. H. New data on the genus *Diadeloplax* Hoare and Mapes, 1995. *J. Paleontol.* **74**, 976–978 (2000).
- Runnegar, B., Pojeta, J., Taylor, M. E. & Collins, D. New species of the Cambrian and Ordovician chitons *Matthevia* and *Chelodes* from Wisconsin and Queensland: evidence for the early history of polyplacophoran mollusks. *J. Paleontol.* **53**, 1374–1394 (1979).
- Sutton, M. D., Briggs, D. E. G., Siveter, David J. & Siveter, Derek J. An exceptionally preserved vermiform mollusk from the Silurian of England. *Nature* **410**, 461–463 (2001).
- Sutton, M. D., Briggs, D. E. G., Siveter, David J. & Siveter, Derek J. Computer reconstruction and analysis of the vermiform mollusk *Acaenoplax hayae* from the Herefordshire Lagerstätte (Silurian, England), and implications for molluscan phylogeny. *Palaeontology* (in the press).
- Haszprunar, G. Is the Aplousobranchia monophyletic? A cladistic point of view. *Am. Malacol. Bull.* **15**, 115–130 (2000).
- Scheltens, A. H. & Ivanov, D. L. An aplousobranch postlarva with iterated dorsal groups of spicules and skeletal similarities to Paleozoic fossils. *Invertebr. Biol.* **121**, 1–10 (2002).
- Conway Morris, S. & Peel, J. S. Articulated halkieriids from the Lower Cambrian of north Greenland. *Nature* **345**, 802–805 (1990).
- Conway Morris, S. & Peel, J. S. Articulated halkieriids from the Lower Cambrian of North Greenland and their role in early protostome evolution. *Phil. Trans. R. Soc. Lond. B* **347**, 305–358 (1995).
- Cohen, B. L., Holmer, L. E. & Lütter, C. The brachiopod fold hypothesis: a neglected body plan hypothesis. *Palaeontology* **46**, 59–65 (2003).
- Williams, A. & Holmer, L. E. Shell structure and inferred growth functions and affinities of the sclerites of the problematic *Micrina*. *Palaeontology* **45**, 845–873 (2002).
- Kniprath, E. Ontogenetic plate and plate field development in two chitons, *Middendorffia* and *Ischnochiton*. *Roux Arch. Dev. Biol.* **189**, 97–106 (1980).
- Van Name, W. G. The supposed Paleozoic barnacle *Protobalanus* and its bearing on the origin and phylogeny of the barnacles. *Am. Mus. Novit.* **197**, 1–8 (1925).
- Aurivillius, C. W. S. Ueber einige Ober-Silurische Cirripeden aus Gotland. *K. Svenska Vetensk. Akad. Handl.* **18**, 1–24 (1892).
- Adrain, J. M. Machaeridian classification. *Alcheringa* **16**, 15–32 (1992).
- Hoare, R. D. & Mapes, R. H. Late Paleozoic problematic sclerites of hercolepadid affinities. *J. Paleontol.* **70**, 341–347 (1996).
- Van Name, W. G. A new specimen of *Protobalanus*, supposed Paleozoic barnacle. *Am. Mus. Novit.* **227**, 1–6 (1926).
- Swofford, D. L. *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)* (Version 4.0b10) (Sinauer Associates, Sunderland, Massachusetts, 2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We are grateful to M. Graham and the staff of the US Army Corps of Engineers, Middle Wabash Area, for their cooperation in the collection and study of this fossil. We thank R. Mapes, R. Hoare, B. Haissini, E. Landing, D. Blake and C. Mah for lending specimens. K. McKeegan, G. Jarzabinski and R. Turner helped with scanning electron microscopy work and G. Gastony with the Greek. The computerized tomography scanning and image processing was carried out by M. Colbert at the X-ray CT Laboratory, University of Texas at Austin, with funding from a US National Science Foundation IGERT Fellowship to T.E.W. C. Fernandez drew the reconstruction and made the epoxy casts of modern chiton aesthete canals. G. Gastony, D. Eernisse and L. Rieseberg provided comments on the manuscript. R. Raff provided advice throughout the project.

Authors' contributions All three authors made significant contributions to this article.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.J.V. (mikev@ess.ucla.edu).

Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic

Matthew M. Mills¹, Celine Ridame¹, Margaret Davey^{2,3}, Julie La Roche¹ & Richard J. Geider²

¹Marine Biogeochemistry, IFM-GEOMAR Leibniz-Institut für Meereswissenschaften, Düsterbrook Weg 20, D-24105 Kiel, Germany

²Department of Biological Sciences, University of Essex, Colchester CO4 3SQ, UK

³Marine Biological Association of the United Kingdom, Plymouth PL1 2PB, UK

The role of iron in enhancing phytoplankton productivity in high nutrient, low chlorophyll oceanic regions was demonstrated first through iron-addition bioassay experiments¹ and subsequently confirmed by large-scale iron fertilization experiments². Iron supply has been hypothesized to limit nitrogen fixation and hence oceanic primary productivity on geological timescales³, providing an alternative to phosphorus as the ultimate limiting nutrient⁴. Oceanographic observations have been interpreted both to confirm and refute this hypothesis^{5,6}, but direct experimental evidence is lacking⁷. We conducted experiments to test this hypothesis during the *Meteor 55* cruise to the tropical North Atlantic. This region is rich in diazotrophs⁸ and strongly impacted by Saharan dust input⁹. Here we show that community primary productivity was nitrogen-limited, and that nitrogen fixation was co-limited by iron and phosphorus. Saharan dust addition stimulated nitrogen fixation, presumably by supplying both iron and phosphorus^{10,11}. Our results support the hypothesis that aeolian mineral dust deposition promotes nitrogen fixation in the eastern tropical North Atlantic.

There is long-standing uncertainty as to whether N or P is the nutrient that limits phytoplankton productivity in the sea¹². On timescales of one or two days, nutrient-enrichment experiments indicate that primary productivity is N-limited in the Sargasso Sea¹³. On geological timescales, however, N₂ fixation can increase the nitrate inventory of the ocean, thus increasing primary production. In turn, N₂ fixation may be limited by either P (ref. 4) or Fe (ref. 3). These two essential nutrients are in sparse supply in oligotrophic oceans, and both P and Fe have been implicated individually as the nutrient that ultimately limits oceanic primary production and carbon dioxide sequestration^{3,4}. However, direct experimental assessment of the relative importance of P and Fe in controlling N₂ fixation in natural plankton populations is lacking⁷.

In the oligotrophic North Atlantic^{5,6,14} N₂ fixation may fuel 50% of the export production¹⁵. Atmospheric transport of Saharan dust is a major source of dissolved Fe to the tropical North Atlantic⁹ and *Trichodesmium* spp. are most abundant in areas of high Saharan dust deposition^{8,16}. In this region, Fe may be supplied in excess of the *Trichodesmium* spp. growth requirements^{5,14} and this has been interpreted to support the P-limitation hypothesis⁵. Furthermore, in the subtropical and tropical Atlantic, the rate of nitrogen fixation has been correlated with the phosphorus content (but not the iron content) of filamentous non-heterocystous *Trichodesmium* spp.⁵. These oceanographic observations have been interpreted to support P-limitation of diazotrophy. In fact, the low dissolved inorganic phosphate concentrations in surface waters of the Sargasso Sea have been used to argue that the phytoplankton community, as a whole, is P-limited in this region⁶.

Nutrient-addition bioassays, designed to investigate which nutrient (N, P or Fe) limits primary production and nitrogen fixation, were conducted at three stations in the tropical Eastern Atlantic (see Supplementary Information) during October–November 2002. A fully replicated, factorial design of the nutrient additions using

trace-metal clean techniques was implemented to assess the individual and combined effects of N, P and Fe on the community as a whole and on the diazotrophs. Nitrogen was supplied as NH₄⁺ and NO₃⁻ to allow for species-specific preferences, common among the picophytoplankton¹⁷. Labelling with ¹⁵N₂ and ¹⁴CO₂ permitted the assessment of nitrogen fixation¹⁸ by diazotrophs and carbon fixation¹⁹ by whole plankton community. Biomass was assessed from chlorophyll *a* concentration²⁰. Additions of Saharan soils were made to assess the potential impact of dust deposition on the primary and diazotrophic production.

At all locations, chlorophyll and nutrient concentrations reflected oligotrophic conditions (see Supplementary Information). Iron concentrations ranged between 1 and 3 nM (P. Croot, personal communication), and are typical of surface waters around these areas¹⁶. Non-heterocystous filamentous cyanobacteria of the genera *Trichodesmium* and *Katagnymene* were the dominant diazotrophs, but based on *nifH* gene sequencing²¹, unicellular cyanobacterial and heterotrophic diazotrophs were also present.

In all experiments, phytoplankton production and biomass were limited by the availability of fixed nitrogen as demonstrated by the 2–3-fold stimulation of CO₂ fixation and the 1.5–2.5-fold increase of chlorophyll concentration in the treatments amended with N. In contrast, additions of P, Fe, or P + Fe did not increase carbon fixation or chlorophyll concentration (Fig. 1). However, once N-limitation was artificially removed by adding nitrate and ammonium, additions of P and Fe further increased productivity and chlorophyll. The largest effect was found when N, P and Fe were added simultaneously, which led to 5–10-fold increases in CO₂ fixation rates (Fig. 1a–c) and corresponding increases in chlorophyll (Fig. 1d–f).

Our experiments demonstrate that N was the proximate limiting nutrient in this region, contrary to recent suggestions of P-limitation throughout the North Atlantic^{6,22}. Nitrogen limitation of primary production accentuates the potential importance of diazotrophy throughout this region. Nitrogen fixation was detected at all stations, but was lowest at the westernmost site (Fig. 1g). Addition of P and Fe together resulted in a 2–3-fold enhancement of the N₂ fixation rate relative to the control at all three stations (Fig. 1g–i). Addition of either P alone or Fe alone did not stimulate N₂ fixation at the two easternmost sites. At these locations, N₂ fixation was co-limited by both P and Fe.

Saharan dust additions stimulated N₂ fixation as much as twofold (Fig. 1g–i), implying that these treatments relieved the co-limitation of N₂ fixation by P and Fe. We calculated that the supply of dissolved P and Fe from the dust additions was sufficient to meet the demands of the enhanced N₂ fixation rates (see Methods and Supplementary Information). We cannot rule out, however, that the alleviation of P-limitation of diazotrophy by the dust addition was due to the addition of another micronutrient, such as Zn, which is a cofactor in many alkaline phosphatases⁷.

The additions of ammonium and nitrate inhibited N₂ fixation (Fig. 1h, i). Diazotrophy may have been suppressed through physiological feedback inhibition of the nitrogenase by dissolved inorganic nitrogen⁸ in these treatments. Alternatively, other microorganisms may have outcompeted the diazotrophs for the limited P and Fe on alleviation of N-limitation. We cannot differentiate between these alternate explanations, but the repression of N₂ fixation by dissolved fixed nitrogen suggests that N-limitation of community primary productivity is a pre-requisite for high rates of diazotrophy in surface waters.

The tropical Atlantic is subjected to some of the highest mineral dust deposition rates in the world⁹, and has high dissolved iron concentrations in surface waters relative to other oceanic basins. As such, this is a region where the phytoplankton community as a whole²³, and the diazotrophs in particular¹⁴, are least likely to be Fe-limited. Given the reported importance of P-limitation in the subtropical and tropical Atlantic^{5,6} and the high Fe concentrations

measured in our study (see Supplementary Information), we were surprised to find that Fe addition was required to stimulate diazotrophy. However, our results are consistent with the recent suggestion that iron may control the abundance of diazotrophs in the South China Sea, another region subject to high dust deposition²⁴. As it is generally argued that Fe concentrations in our study area are in excess of diazotroph iron requirements^{14,25}, our findings suggest that total dissolved Fe concentration is a poor index of bioavailability²⁶, perhaps due to temporal variation in the chemical speciation of dissolved Fe (ref. 7). It is also possible that the level of iron required to saturate diazotroph growth has been underestimated. The important role of Fe in our study region implies that the control of N_2 fixation by Fe should be even greater in other oceanic regions^{6,14} that receive less dust deposition.

Whether P and Fe co-limitation of N_2 fixation is peculiar to the eastern tropical North Atlantic during the time that we sampled, or is common throughout the oligotrophic ocean, can only be ascertained by further research. It is usually assumed that P is supplied primarily from deep waters by mixing or diffusion, whereas Fe is supplied both from deep waters and by mineral dust deposition from the atmosphere. Temporal variability in these sources of nutrients may cause transients in the ratio of P supply to Fe supply, and thus affect the relative importance of these two elements in controlling N_2 fixation. Previous suggestions of P-limitation of diazotrophy in the North Atlantic^{5,6} were based on observations made during spring when dust deposition is at its highest. Although dust deposition in the tropical Atlantic is high relative to other locations, the seasonal low deposition during our study in autumn may have shifted the balance towards P and Fe co-limitation. (see Supplementary Information).

A common source of P and Fe, through dust deposition, could also lead to co-limitation of nitrogen fixation. Our observation of stimulation of nitrogen fixation by the addition of Saharan dust is consistent with this suggestion. To be valid, this hypothesis requires

that aeolian supply of the bioavailable forms of P and Fe must roughly match the diazotrophic demand for P and Fe. It is commonly assumed that the dissolution of dust can supply the Fe required to support nitrogen fixation, but would not be able to supply the necessary P. This is because crustal material contains about 30 times as much Fe as P (3.5% versus 0.11%), whereas diazotrophs require 30–300 times less Fe than P. However, there are uncertainties on both the supply and demand sides of this comparison. The availability of these nutrients on contact with sea water depends on the dust source, the dust load to the surface of the ocean, prevailing pH conditions during atmospheric processing as well as the mode of deposition (wet or dry)¹¹. Due to solubility differences, more phosphate than Fe can be released from dust (see Supplementary Information). Furthermore, we expect the dissolved phosphate to be immediately bioavailable. The bioavailability of the Fe released from dust is less certain²⁶. On the demand side, estimates of P and Fe requirements of *Trichodesmium*^{14,27} vary, whereas those of the recently discovered marine diazotrophs²¹ are not yet characterized. The possibility that simultaneous release of P and Fe from dust may stimulate diazotroph production in regions of high dust deposition remains an open question (see Supplementary Information).

The results presented here have important implications for understanding controls on marine N_2 fixation and how it relates to CO_2 fixation in the North Atlantic. First, contrary to recent suggestions^{6,22}, our experiments demonstrate that the total primary productivity of the natural plankton community in the tropical Atlantic is N-limited. Second, they demonstrate that N_2 fixation is co-limited by Fe and P in a region where mineral dust deposition is high and iron should be in excess. This has not been reported before. Further studies are required to determine whether this co-limitation is widespread. Finally, our results suggest that dust, when supplied at high levels locally, can relieve Fe and P co-limitation of diazotrophy. The tropical North Atlantic is a region of high dust deposition. It is also considered one of the most important areas

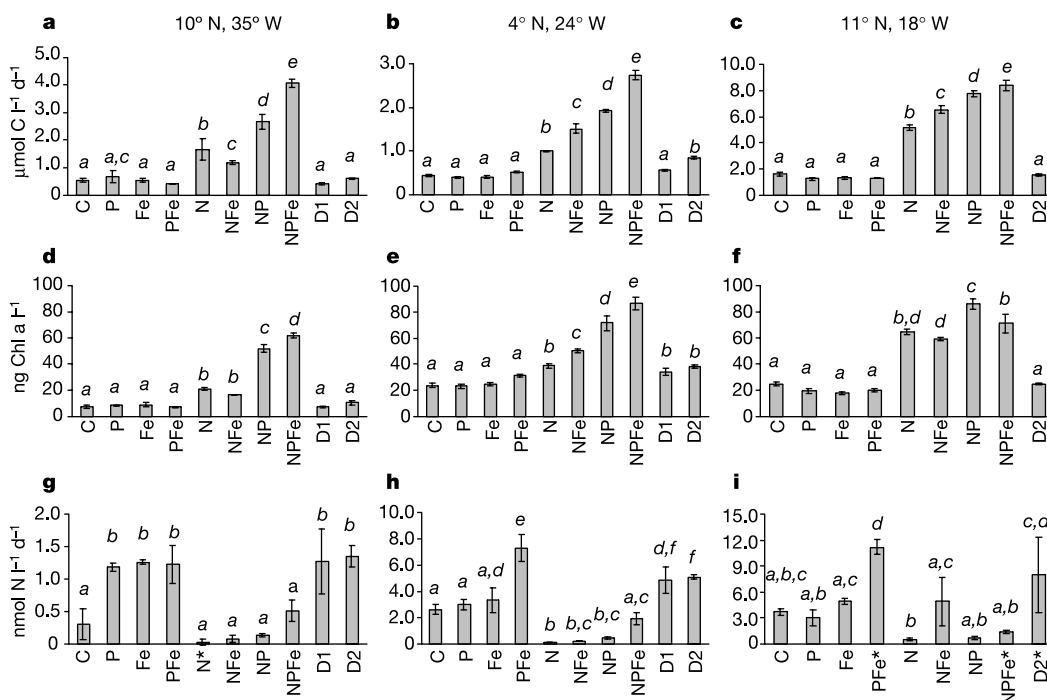


Figure 1 Effect of nutrient additions during bioassay experiments. Measurements were taken at three sites in the tropical Atlantic during October–November 2002. **a–c**, Net CO_2 fixation rates; **d–f**, chlorophyll *a* concentration; **g–i**, N_2 fixation rates. Carbon fixation, nitrogen fixation and chlorophyll were measured from separate triplicate bottles, such that nine bottles were incubated for each nutrient treatment. Shown are means $\pm$ standard

errors, $n = 3$ for all variables except where $n = 2$ (as indicated by an asterisk). Treatment means were compared using a one-way ANOVA and a Fisher PLSD means comparison test. Means that are not significantly different are labelled with the same letter ($\alpha = 0.05$).

globally for N_2 fixation¹⁵. Dust deposition at this location is highly episodic, and has varied widely on geological timescales²⁸. If dust deposition can to some extent relieve both P and Fe limitation of diazotrophy, the postulated link between climate-driven changes in dust deposition and N_2 fixation may be even stronger than initially suggested^{3,29}. □

Methods

Trace-metal clean techniques were strictly used throughout the preparation and execution of the experiments as this is crucial to the good survival of diazotrophs and for *Trichodesmium* spp. in particular²⁷. Surface sea water was collected (1–3 m) after dark using a trace-metal clean diaphragm pump. Sea water was pumped into 60-l carboys from which it was siphoned into 1.8 l acid-washed polycarbonate bottles. Under a laminar flow hood, nutrients were added alone and in combination to final concentrations of $1.0 \mu M NH_4^+$ + $1.0 \mu M NO_3^-$, $0.2 \mu M NaH_2PO_4$, and $2.0 nM FeCl_3$. Saharan dust treatments were also conducted with final concentrations of $0.5 mg l^{-1}$ (D1) and $2 mg l^{-1}$ (D2). These concentrations were chosen to simulate concentrations in the upper 1 m of the water column after a strong Saharan aerosol deposition event¹¹. The dust consisted of the fine fractions of surface soils collected in the Hoggar region (Southern Algeria) with a grain-size distribution and chemical composition typical for Saharan aerosols collected far from the source^{10,11}. The measured P and Fe content of the dust was $0.14 \pm 0.01\%$ and $4.97 \pm 0.49\%$ ($\pm$ one standard error). The phosphate liberated from the dust treatments was approximately 2.7 and $10.8 nmol l^{-1}$, and Fe released was 0.9 and $3.6 nmol l^{-1}$, for the 0.5 and $2.0 mg l^{-1}$ dust treatments respectively (see Supplementary Information). The bottles were then sealed gas tight and placed in an on-deck incubator with circulating surface sea water. Light was attenuated to 20% of incident surface values with blue filters (Lagoon Blue, Lee Filters #172). For each treatment, parallel incubations for each variable (carbon fixation, nitrogen fixation and biomass) were run in triplicate over 48 h with rate measurements made during the final 24 h and chlorophyll concentration determined at 48 h. At each of the three study sites, over 100 bottles were incubated. For measuring net nitrogen fixation rates, $1.0 ml$ of $99\% ^{15}N_2$ was introduced to each bottle through a butyl septum using a gas-tight syringe. $^{15}N_2$ uptake measurements may underestimate nitrogen fixation if significant release of dissolved nitrogen occurs³⁰. However, this effect should be minimal in our experiments because, in a N-limited oligotrophic system, our 24-h rate measurements should allow released labile dissolved N to be reincorporated into particulate matter. For measuring primary productivity, $0.1 mCi ^{14}C$ -bicarbonate was added to each bottle. All incubations were conducted from dawn-to-dawn and stopped by gentle filtration.

Received 17 September 2003; accepted 5 April 2004; doi:10.1038/nature02550.

- Martin, J. H. & Fitzwater, S. E. Iron deficiency limits phytoplankton growth in the north-east Pacific subarctic. *Nature* **331**, 341–343 (1988).
- Boyd, P. *et al.* A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilization. *Nature* **407**, 695–702 (2000).
- Falkowski, P. G. Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean. *Nature* **387**, 272–275 (1997).
- Tyrrell, T. The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature* **400**, 525–531 (1999).
- Sanudo-Wilhelmy, S. A. *et al.* Phosphorus limitation of nitrogen fixation by *Trichodesmium* in the central Atlantic Ocean. *Nature* **411**, 66–69 (2001).
- Wu, J., Sunda, W., Boyle, E. A. & Karl, D. M. Phosphate depletion in the Western North Atlantic Ocean. *Science* **289**, 759–762 (2000).
- Morel, F. M. M. & Price, N. M. The biogeochemical cycles of trace metals in the oceans. *Science* **300**, 944–947 (2003).
- Capone, D. G., Zehr, J. P., Paerl, H. W., Bergman, B. & Carpenter, E. J. *Trichodesmium*, a globally significant cyanobacterium. *Science* **276**, 1221–1229 (1997).
- Gao, Y., Kaufman, Y. J., Tanre, D., Kolber, D. & Falkowski, P. G. Seasonal distributions of aeolian iron fluxes to the global ocean. *Geophys. Res. Lett.* **28**, 29–32 (2001).
- Guieu, C., Loye-Pilot, M. D., Ridame, C. & Thomas, C. Chemical characterization of the Saharan dust end-member: Some biogeochemical implications for the western Mediterranean Sea. *J. Geophys. Res. Atmos.* **107**, 4258 (2002).
- Ridame, C. & Guieu, C. Saharan input of phosphate to the oligotrophic water of the open western Mediterranean Sea. *Limnol. Oceanogr.* **47**, 856–869 (2002).
- Redfield, A. C. The biological control of chemical factors in the environment. *Am. Sci.* **46**, 205–221 (1958).
- Graziano, L. M., Geider, R. J., Li, W. K. W. & Olaiola, M. Nitrogen limitation of North Atlantic phytoplankton: Analysis of physiological condition in nutrient enrichment experiments. *Aquat. Microb. Ecol.* **11**, 53–64 (1996).
- Berman-Frank, I., Cullen, J. T., Shaked, Y., Sherrell, R. M. & Falkowski, P. G. Iron availability, cellular iron quotas, and nitrogen fixation in *Trichodesmium*. *Limnol. Oceanogr.* **46**, 1249–1260 (2001).
- Gruber, N. & Sarmiento, J. L. Global patterns of marine nitrogen fixation and denitrification. *Glob. Biogeochem. Cycles* **11**, 235–266 (1997).
- Tyrrell, T. *et al.* Large-scale latitudinal distribution of *Trichodesmium* spp. in the Atlantic Ocean. *J. Plankton Res.* **25**, 405–416 (2003).
- Palinska, K. A. *et al.* The signal transducer P-II and bicarbonate acquisition in *Prochlorococcus marinus* PCC 9511, a marine cyanobacterium naturally deficient in nitrate and nitrite assimilation. *Microbiology* **148**, 2405–2412 (2002).
- Montoya, J. P., Voss, M., Kähler, P. & Capone, D. G. A simple, high-precision, high-sensitivity tracer assay for N_2 fixation. *Appl. Environ. Microbiol.* **62**, 986–993 (1996).
- Lohrenz, S. *et al.* A comparison of *in situ* and simulated *in situ* methods for estimating primary production. *J. Plankton Res.* **14**, 201–221 (1992).
- Welschmeyer, N. Fluorometric analysis of chlorophyll *a* in the presence of chlorophyll *b* and

- pheopigments. *Limnol. Oceanogr.* **38**, 1985–1992 (1994).
- Zehr, J. P. *et al.* Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. *Nature* **412**, 635–638 (2001).
 - Ammerman, J. W., Hood, R. R., Case, D. A. & Cotner, J. B. Phosphorus deficiency in the Atlantic: An emerging paradigm in oceanography. *EOS Trans. Am. Geophys. Union* **84**, 165–170 (2003).
 - Coale, K. H., Fitzwater, S. E., Gordon, R. M., Johnson, K. S. & Barber, R. Control of community growth and export production by upwelled iron in the equatorial Pacific Ocean. *Nature* **379**, 621–624 (1996).
 - Wu, J. F. *et al.* Dissolved inorganic phosphorus, dissolved iron, and *Trichodesmium* in the oligotrophic South China Sea. *Global Biogeochem. Cycles* **17**, 1008 (2003).
 - Kustka, A., Sañudo-Wilhelmy, S., Carpenter, E. J., Capone, D. G. & Raven, J. A. A revised estimate of the iron use efficiency of nitrogen fixation, with special reference to the marine cyanobacterium *Trichodesmium* spp. (Cyanophyta). *J. Phycol.* **39**, 12–25 (2003).
 - Rue, E. & Bruland, K. Complexation of Fe(III) by natural ligands in the central north Pacific as determined by a new competitive ligand equilibrium/absorptive cathodic voltammetry method. *Mar. Chem.* **50**, 117–138 (1995).
 - Kustka, A. *et al.* Iron requirements for dinitrogen- and ammonium-supported growth in cultures of *Trichodesmium* (IMS 101): Comparison with nitrogen fixation rates and iron: carbon ratios of field populations. *Limnol. Oceanogr.* **48**, 1869–1884 (2003).
 - Mahowald, N. *et al.* Dust sources and deposition during the last glacial maximum and current climate: A comparison of model results with paleodata from ice cores and marine sediments. *J. Geophys. Res. Atmos.* **104**, 15895–15916 (1999).
 - Karl, D. *et al.* Dinitrogen fixation in the world's oceans. *Biogeochemistry* **57**, 47–98 (2002).
 - Glibert, P. M. & Bronk, D. A. Release of dissolved organic nitrogen by marine diazotrophic cyanobacteria, *Trichodesmium* spp. *Appl. Environ. Microbiol.* **60**, 3996–4000 (1994).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We would like to acknowledge the captain and crew of the F/S *Meteor* and the scientists aboard the *Meteor 55* SOLAS cruise, especially D. Wallace, K. Lochte, H. Bange, P. Croot, M. Voss, F. Malien, R. Langlois and P. Fritsche. We thank D. Wallace for insightful comments on this manuscript. Additionally, we acknowledge W. Balzer for loaning the clean container used during M55. This work was supported by the Deutsche Forschungsgemeinschaft's Meteor Schwerpunktprogramm, a Natural Environment Research Council grant to R.J.G., and a Marie Curie Post Doctoral fellowship to C.R.

Authors' contributions The first three authors made equal contributions to the success of the experiments. This manuscript is the product of an equal collaboration between the groups of J.L. at IfM-Geomar and R.J.G. at University of Essex.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.J.G. (geider@essex.ac.uk).

Evidence for ecology's role in speciation

Jeffrey S. McKinnon¹, Seiichi Mori², Benjamin K. Blackman^{3*}, Lior David^{3*}, David M. Kingsley³, Leia Jamieson¹, Jennifer Chou¹ & Dolph Schluter⁴

¹Biological Sciences, University of Wisconsin-Whitewater, Whitewater, Wisconsin 53190, USA

²Biological Laboratory, Gifu-Keizai University, Ogaki, Gifu prefecture 503-8550, Japan

³HHMI and Stanford University School of Medicine, Beckman Center B300, 279 Campus Drive, Stanford, California 94305-5329, USA

⁴Department of Zoology and Centre for Biodiversity, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

* Present addresses: Department of Biology, Indiana University, Jordan Hall Rm 325, 1001 E. 3rd Street, Bloomington, Indiana 47405, USA (B.K.B.); Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA (L.D.)

A principal challenge in testing the role of natural selection in speciation is to connect the build-up of reproductive isolation between populations to divergence of ecologically important traits^{1,2}. Demonstrations of 'parallel speciation', or assortative mating by selective environment, link ecology and isolation^{3–5}, but the phenotypic traits mediating isolation have not been confirmed. Here we show that the parallel build-up of mating

incompatibilities between stickleback populations can be largely accounted for by assortative mating based on one trait, body size, which evolves predictably according to environment. In addition to documenting the influence of body size on reproductive isolation for stickleback populations spread across the Northern Hemisphere, we have confirmed its importance through a new experimental manipulation. Together, these results suggest that speciation may arise largely as a by-product of ecological differences and divergent selection on a small number of phenotypic traits.

Anadromous and stream-resident threespine stickleback (*Gasterosteus aculeatus*) populations that presently breed sympatrically often show little introgression and exhibit pre-mating isolation in the laboratory^{6,7} (but see ref. 8). Stream populations from different regions are phenotypically similar and the stream ecotype has evolved repeatedly, whereas the ancestral, anadromous ecotype has persisted in similar but geographically widespread marine habitats^{6,9}. The repeated origin of like stream forms and the maintenance of the large anadromous phenotype in distant but similar habitats suggest that certain morphological traits, including body size, are principally adaptations to environmental selection regime^{6,10}, a view supported by experimental studies of other stickleback systems^{6,7,11,12}. The large size of anadromous sticklebacks may be an adaptation to relatively long migrations¹³ but other selective factors such as predation regime and food supply might also contribute. Differences in size between stickleback populations have been shown to be substantially heritable in common-garden studies^{7,14}, including one of anadromous and stream populations from California¹³ and a recent investigation that identifies a major quantitative trait locus for size distinguishing a Japanese marine and a British Columbia freshwater population¹⁵.

The wide distribution of stream and anadromous sticklebacks provides exceptional replication for investigation of the factors causing reproductive isolation^{6,7}. We collected subjects for mating trials from geographically distant regions (Alaska, British Columbia, Iceland (stream only), Scotland, Norway (stream only) and

Japan; details in Supplementary Information) chosen to maximize the probability that the stream populations had evolved independently and that the anadromous populations were physically isolated from one another, with minimal gene flow between them. Thus we focused primarily on sites that had been recently glaciated and were separated from other study locales by large expanses of land, sea or both. Nested AMOVA's (analysis of molecular variance) based on microsatellite data support our assumption of close relationships among geographically adjacent populations rather than those of the same ecotype (Table 1). These analyses consistently reveal that much of the variation in microsatellites is accounted for by geography, as expected if stream populations have evolved repeatedly from anadromous populations inhabiting the same region. Ecotype should account for a significant proportion of molecular variance if the stream ecotype evolved just once then spread to the various study sites, but this is never observed. Results from analyses of allozymes¹⁶ and mitochondrial DNA sequences¹⁷ also suggest replicated origins of freshwater populations from anadromous or marine ancestors, and distant relationships between far-flung anadromous populations. Introgression within regions and insufficient time for full lineage sorting may contribute to these patterns; however, the glacial history of most of our study areas and the well documented rapid pace of stickleback evolution¹⁸ support the hypothesis that the stream-resident ecotype has evolved repeatedly.

Consistent with earlier reports^{6,10}, anadromous populations in our study possessed larger average body sizes than the stream populations (mean anadromous female length = 69.2 mm, standard error (SE) = 5.3 mm; mean stream female standard length = 46.6 mm, SE = 1.6 mm; $P = 0.001$, two-tailed t -test, $n = 10$ populations; a paired test for regions with both ecotypes was also significant). To assess reproductive compatibility or isolation, recently collected fish from all sites were brought together in the laboratory for mating tests.

Nearby stream and anadromous pairs exhibited overall high, significant isolation. Courtship was more than twice as successful between pairs of the same ecotype, and in this case from the same population, relative to pairs composed of different ecotype individuals (Fig. 1; see also refs 19, 20).

A significant pattern of parallel speciation, more precisely parallel reproductive isolation, emerged in tests of pairs composed exclusively of individuals from different regions (Fig. 1). Despite independent origins and/or great distances between populations, female stream and anadromous sticklebacks each preferred males of their own ecotype. This pattern provides a strong signature of divergent

Table 1 Results of AMOVA

Grouping model	Percentage of variation	F_{ST}/ϕ_{ST}
Continent (all populations)		
Among continents	7.85/20.90	0.08†/0.21‡
Among populations in continent	14.93/17.54	0.16†/0.22‡
Within populations	77.22/61.56	0.23‡/0.38‡
Ocean (all populations)		
Atlantic versus Pacific	6.13/13.35	0.06†/0.13†
Among populations in ocean	17.13/25.23	0.18‡/0.29‡
Within populations	76.74/61.42	0.23‡/0.39‡
Ecotype (all populations)		
Anadromous versus stream	-0.07/-0.91	-0.01/-0.01
Among populations in ecotype	21.04/35.13	0.21†/0.34‡
Within populations	79.04/65.79	0.21†/0.34‡
Region (single ecotype regions excluded)		
Among regions	2.32/12.12	0.02*/0.12†
Among populations in regions	19.65/26.20	0.20‡/0.29‡
Within populations	78.03/61.69	0.21†/0.38‡
Ecotype (single ecotype regions excluded)		
Anadromous versus stream	-0.80/-2.61	-0.01/-0.01
Among populations in ecotype	22.24/39.13	0.22†/0.38‡
Within populations	78.56/63.48	0.21†/0.36‡

Values to the left of the solidus represent F_{ST} -based analyses whereas values to the right of the solidus represent ϕ_{ST} -based analyses (all conducted with Arlequin, as were permutation tests of significance). The former incorporates allele frequency variation only; the latter incorporates variation in allele size and frequency. Continents are Europe (Iceland, Norway, Scotland), Japan (that is, Asia) and North America (Alaska, British Columbia). Regions are defined as groups of populations from Alaska, British Columbia, Iceland, Japan, Norway and Scotland, but only regions for which data are available both from anadromous and stream-resident populations are included in these analyses of region; thus Norway and Iceland stream populations are excluded. Estimates of divergence time between populations, based on $\delta\mu^2$, range from less than 5,000 to over 200,000 years and are given in Supplementary Information along with additional phylogenetic analyses and tables of genetic distances.

* $P < 0.1$.

† $P < 0.01$.

‡ $P < 0.001$.

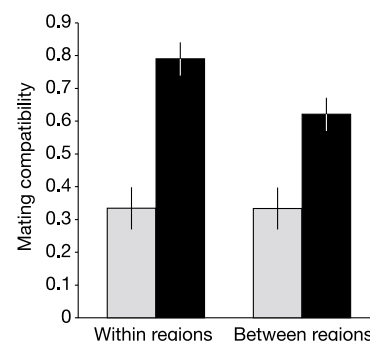


Figure 1 Mating compatibility (mean proportion of trials involving a nest inspection) for same (black) and different ecotype (grey) combinations, for within region tests ($P = 0.0058$, paired one-tailed t -test, $n = 7$ populations) and between region tests ($P = 0.0045$, $n = 9$; when the genetic distance measure $\delta\mu^2$ and ecotype match were analysed together, only ecotype was significant: $\delta\mu^2$, $P = 0.083$; ecotype match, $P = 0.0036$, one-tailed t -tests, $n = 9$). Total $N = 850$ (see Supplementary Information). Error bars are 1 SE.

selection³, suggesting that reproductive isolation is brought about by adaptation to different environments. Moreover, these patterns stand out at vast distances, showing that the influence of ecology and selection is sufficient to be detectable even above the background variation that must exist on such a scale.

Levels of reproductive isolation are well accounted for by differences in a single trait, body size. Stream sticklebacks are typically smaller than anadromous fish and the probability of mating between populations is negatively related to size difference, even after accounting for the time that populations have been separated (Fig. 2). When body size difference and ecotype match are considered together, only size difference is significant ($P = 0.014$; for ecotype, $P = 0.10$, one-tailed t -tests, $n = 9$ in both cases). Thus size-assortative mating seems to be a very general tendency, with mating preferences and patterns that cause reproductive isolation changing largely as by-products of differences in body size. Nevertheless, the near significance of the term for residual ecotype match raises the possibility that traits other than body size make a secondary contribution.

Experimental manipulation of body size directly confirms the connection between size divergence and the build-up of reproductive isolation. We raised large and small females of each ecotype from both Japan and British Columbia, mainly by providing long and short growing periods, and tested mating patterns using field-collected British Columbia males. As predicted, the preferred male ecotype depended on the size to which the female was raised (Fig. 3). Usually isolated, different ecotype combinations were compatible if size differences were small as a result of the manipulation; for example, experimentally induced small British Columbia anadromous females and naturally small field-caught stream males courted with considerable success (mean standard length difference = 3.5 mm, mating compatibility = 0.5). Conversely, same-ecotype pairings were relatively unsuccessful when manipulation led to large size differences; for example, mating compatibility was just 0.25 for experimentally induced small British Columbia anadromous females with large field-caught anadromous males (mean standard length difference = 17.5 mm, Fig. 3). This mating pattern did not vary significantly with female ecotype or

region. Females also retained a statistically independent and significant, albeit weaker, preference for males of their own ecotype (Fig. 3), once again suggesting that traits other than size do make a secondary contribution to reproductive isolation.

Many of the populations in our comparative analyses are completely allopatric to one another, suggesting that interactions in sympatry, such as through reinforcement^{21,22}, are not essential for the evolution of reproductive isolation. Indeed, populations sympatric with the opposite ecotype did not show stronger preferences for their own population than did allopatric populations not sympatric with the opposite ecotype (same ecotype – different ecotype mean mating compatibility, sympatric populations versus allopatric: $P = 0.433$, one-tailed t -test, $n = 6$), although there was considerable variation in levels of isolation among pairs⁸. Similar results, in terms of preference for the same ecotype, were obtained in the between-region tests for populations that do or do not co-occur with the opposite ecotype ($P = 0.424$, one-tailed t -test, $n = 9$). The pattern of reproductive character displacement may not be observed here, in contrast with some other studies of sticklebacks^{6,21}, because it is masked by the large size differences in many allopatric comparisons and because we did not conduct paired tests of sympatric and allopatric populations within lineages, which is a more sensitive assay.

In contrast to body size, divergence in red colouration does not appear to be an important cause of reproductive isolation in the stream–anadromous stickleback system. Interpopulation assortative mating based on male colouration was not significant for these populations (for allopatric population combinations, $P = 0.76$, one-tailed t -test, $n = 9$ populations). This finding may differ from results for lake populations because of less extreme colour differences in this system than between lake sticklebacks²³.

Our comparative analyses and experimental manipulation together demonstrate the importance of a single phenotypic trait, body size, in stickleback speciation and suggest that divergent selection on this trait makes a predominant contribution to reproductive isolation. Although size has previously been shown to be important in isolation between pairs of stickleback populations^{1,24}, it has not been shown to account for patterns, including parallel

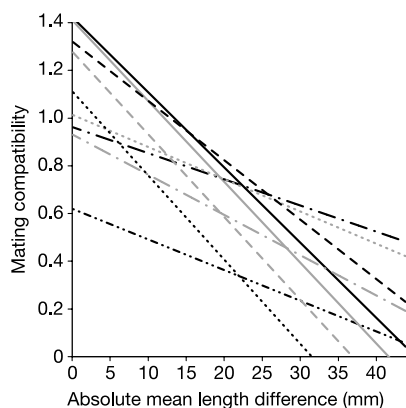


Figure 2 Regression lines for mating compatibility (arcsine square-root-transformed) on absolute mean standard length difference for each female population (Supplementary Information) tested with allopatric male populations. The decline of mating compatibility with increasing difference in body size is highly significant ($P < 0.0001$, one-tailed t -test, $n = 9$). The relationship between genetic distance/time ($\delta\mu^2$) and reproductive isolation was not significant ($P = 0.10$) and remained so when analysed with size difference ($P = 0.070$), whereas size continued to be significant ($P < 0.0001$). Statistical significance of these results is unchanged in Mantel tests. Regressions of mating compatibility on Nei's standard distance and (1-proportion shared alleles) were each just significant ($P = 0.035$ and 0.018 , respectively) but did not approach significance in joint analyses with size difference ($P > 0.13$ in both cases).

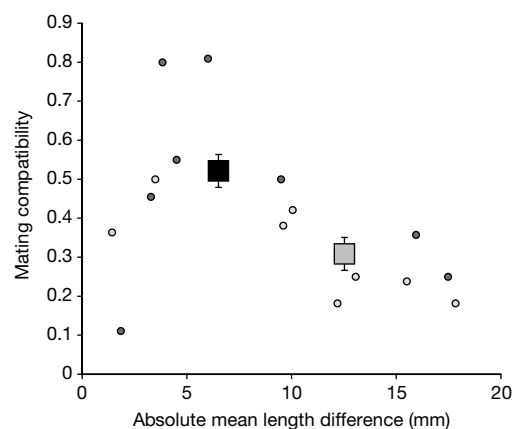


Figure 3 Mating compatibility versus absolute mean standard length difference between males and females, with female size manipulated. Circles represent each combination of female region, ecotype, size manipulation and male ecotype, with same ecotype pairings coloured dark grey. Squares are mean values for pairs manipulated to be similar (black: large (size-manipulated) females with anadromous males, which are large, and small females with stream males, which are small) or different (grey: large females with stream males and small females with anadromous males) in size ($N = 262$; error bars are 1 SE). The preferred male ecotype depended on the size to which females were manipulated ($\chi^2 = 13.40$, $P = 0.0003$, degrees of freedom = 1, 251) although females also retained a preference for males of their own ecotype ($\chi^2 = 7.46$, $P = 0.0063$, degrees of freedom = 1, 251).

isolation, among sets of populations; nor has its importance been confirmed previously through a manipulation. The comparative patterns documented here were obtained with populations separated, in some cases, for hundreds of thousands of years, extending the role of divergent selection and ecology well beyond that implied by previous work^{4,5}.

The malleability of body size raises the possibility that multiple mechanisms underlie reproductive isolation between populations differing in size; this topic deserves further investigation. In addition to directly heritable effects, phenotypic plasticity may contribute to the large size of anadromous fish through rapid growth in the relatively productive temperate marine habitat. In the extreme case, this mechanism might provide an alternative explanation for our results, if it were the main cause of size differences between some stream and anadromous populations. But whether directly heritable or a plastic by-product of evolved migratory differences and habitat choice, stream–anadromous size differences are essentially ecological in origin and can yield reproductive isolation only when coupled with a widespread tendency towards size-assortative mating. □

Methods

Collection and maintenance of fish

Sticklebacks were collected in the spring of each year using minnow traps and hand seines. They were transported to Vancouver (1996) or Whitewater (2000, 2001) and maintained using standard procedures²⁵.

Microsatellites and genetic distances

Microsatellites were genotyped²⁶ and used to calculate $\delta\mu^2$, the appropriate genetic distance measure when timeframes allow for mutation. We excluded loci violating the stepwise mutation assumption, leaving 15 microsatellites from 22 fish per population. We conducted supplemental analyses using two other genetic distances.

Mating trials

Protocols in 30-min ‘no choice’ trials, conducted with a nesting male and a single female in a 96-l aquarium, were essentially identical to those described in previous studies^{1,4}. Before a male's first trial (2000 and 2001), we scored his red colouration rank visually²³.

We used nest inspection as our principal measure of mating compatibility because females from some populations exhibited very low levels of nest entry with all male populations. Furthermore, nest inspection may be a more robust measure of preference than nest entry, which is more influenced by changes in female state²⁷.

To ensure statistical independence of observations, we used female population as the unit of replication in most analyses, following ref. 4. Tests of assortative mating used paired *t*-tests to compare average frequency of female nest inspection with males of ‘same ecotype’ and ‘other ecotype’ (ecotype match). In tests of assortative mating within a region, ‘same ecotype’ refers to males from the female's own population, whereas ‘other ecotype’ refers to males from the adjacent population but of the opposite life history. In the main, conservative test of parallel reproductive isolation, ‘own ecotype’ refers to males with the same life history as females but from other regions, whereas ‘other ecotype’ includes all males from other regions and with the opposite life history (following refs. 3, 4).

Effects of body size, genetic distance and colouration on mating compatibility were tested with one-sample *t*-tests. The single measurement for each female population was the slope of a linear regression of nest inspection frequency against genetic distance from each male population, or against absolute difference in body size or male colouration, weighting by sample size. The latter analyses excluded data for sympatric populations to avoid any direct influence of familiarity or coevolution. Multiple regression analyses were used to assess independent variables simultaneously, and mean slopes again compared. Ecological similarity, critical to tests of parallel isolation, was included in some analyses using the dummy variable ‘ecotype match’. The arcsine transformation was applied to proportions. Standard and partial Mantel tests were used to check regression results using NTSyspc V 2.11. We focus on regression results owing to possible problems with partial Mantel tests²⁸ but our main findings are statistically robust.

In the central analyses of parallel reproductive isolation, the observed distribution of genetic distances accounts less well for mating patterns, in terms of shared evolutionary history, than a star phylogeny possessing no evolutionary structure (unrooted tree with all internal branches set to 0, tip branches to 1 (ref. 29); shared evolutionary history was calculated from a tree based on $\delta\mu^2$ and constructed using PHYLIP's ‘Kitsch’ v. 3.6a3; best log-likelihood for star phylogeny = 6.32, for Fitch–Margoliash = 4.26). Consequently, we have made no additional statistical adjustments for shared history beyond the incorporation of $\delta\mu^2$ into some analyses.

Size manipulations

Crosses were made and sticklebacks initially raised using standard procedures³⁰. At approximately 8 weeks of age, all juvenile sticklebacks were culled to a maximum density of 45 individuals per 96-l aquarium. Fish were raised at these high densities to one year of age to produce a small body size. Those fish destined to be raised to two years of age and

large size were further reduced to a maximum density of 25 fish per aquarium for their second year; densities were sometimes lower and variable, however, due mainly to mortality. Fifty separate crosses were made and an average of 5.2 females tested from each cross for a total of 262 trials, all conducted with wild-caught British Columbia stream and anadromous males.

We analysed the resulting data with a logistic regression on nest inspection that included the effects female size treatment, female region, female ecotype, male ecotype and their interactions. The key experimental prediction was that female sticklebacks manipulated to large size would be relatively more compatible with anadromous males, which are large, whereas fish manipulated to small size would be more compatible with stream males, which are small. Statistically, this should be manifest in a significant interaction between female size treatment and male ecotype. Assortative mating by ecotype would contribute to a significant male ecotype by female ecotype interaction term. Three-way interactions and the four-way interaction were nonsignificant and omitted from the main analysis. Additional details, including the combinations of populations tested in the comparative experiment, are given in Supplementary Information. Experiments were approved by the UW-Whitewater animal care committee (IACUC).

Received 21 January; accepted 6 April 2004; doi:10.1038/nature02556.

- Nagel, L. & Schluter, D. Body size, natural selection, and speciation in sticklebacks. *Evolution* **52**, 209–218 (1998).
- Greenberg, A. J., Moran, J. R., Coyne, J. A. & Wu, C. I. Ecological adaptation during incipient speciation revealed by precise gene replacement. *Science* **302**, 1754–1757 (2003).
- Schluter, D. & Nagel, L. M. Parallel speciation by natural selection. *Am. Nat.* **146**, 292–301 (1995).
- Rundle, H. D., Nagel, L., Boughman, J. & Schluter, D. Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**, 306–308 (2000).
- Nosil, P., Crespi, B. J. & Sandoval, C. P. Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature* **417**, 440–443 (2002).
- McPhail, J. D. in *The Evolutionary Biology of the Threespine Stickleback* (eds Bell, M. A. & Foster, S. A.) 399–437 (Oxford Univ. Press, Oxford, 1994).
- McKinnon, J. S. & Rundle, H. Speciation in nature: the threespine stickleback model systems. *Trends Ecol. Evol.* **17**, 480–488 (2002).
- Hagen, D. W. Isolating mechanisms in threespine sticklebacks (*Gasterosteus*). *J. Fish. Res. B. Can.* **24**, 1637–1692 (1967).
- Bell, M. A. & Foster, S. A. in *The Evolutionary Biology of the Threespine Stickleback* (eds Bell, M. A. & Foster, S. A.) 1–27 (Oxford Univ. Press, Oxford, 1994).
- Baker, J. A. in *The Evolutionary Biology of the Threespine Stickleback* (eds Bell, M. A. & Foster, S. A.) 144–187 (Oxford Univ. Press, Oxford, 1994).
- Hendry, A. P., Taylor, E. B. & McPhail, J. D. Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. *Evolution* **56**, 1199–1216 (2002).
- Rundle, H. D., Vamori, S. M. & Schluter, D. Experimental test of predation's effect on divergent selection during character displacement in sticklebacks. *Proc. Natl Acad. Sci. USA* **100**, 14943–14948 (2003).
- Snyder, R. J. & Dingle, H. Adaptive, genetically based differences in life history between estuary and freshwater threespine sticklebacks (*Gasterosteus aculeatus* L.). *Can. J. Zool.* **67**, 2448–2454 (1989).
- McPhail, J. D. Inherited interpopulation differences in size at first reproduction in threespine stickleback, *Gasterosteus aculeatus* L. *Heredity* **38**, 53–60 (1977).
- Colosimo, P. F. et al. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol.* **2**, e109 (2004).
- Buth, D. G. & Haglund, T. R. in *The Evolutionary Biology of the Threespine Stickleback* (eds Bell, M. A. & Foster, S. A.) 61–84 (Oxford Univ. Press, Oxford, 1994).
- Orti, G., Bell, M. A., Reimchen, T. E. & Meyer, A. Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent colonizations. *Evolution* **48**, 608–622 (1994).
- Bell, M. A. Lateral plate evolution in the threespine stickleback: getting nowhere fast. *Genetica* **112–113**, 445–461 (2001).
- Hay, D. E. & McPhail, J. D. Mate selection in three-spine sticklebacks (*Gasterosteus*). *Can. J. Zool.* **53**, 441–450 (1975).
- Ziuganov, V. V. Reproductive isolation among lateral plate phenotypes (low, partial, complete) of the threespine stickleback, *Gasterosteus aculeatus*, from the White Sea basin and the Kamchatka Peninsula, Russia. *Behaviour* **132**, 1173–1181 (1995).
- Rundle, H. D. & Schluter, D. Reinforcement of stickleback mate preferences: sympatry breeds contempt. *Evolution* **52**, 200–208 (1998).
- Servedio, M. R. Beyond reinforcement: the evolution of premating isolation by direct selection on preferences and post-mating, prezygotic incompatibilities. *Evolution* **55**, 1909–1920 (2001).
- Boughman, J. W. Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature* **411**, 944–948 (2001).
- Ishikawa, M. & Mori, S. Mating success and male courtship behaviors in three populations of the threespine stickleback. *Behaviour* **137**, 1065–1080 (2000).
- McKinnon, J. S. Video mate preferences of female three-spined sticklebacks from populations with divergent male coloration. *Anim. Behav.* **50**, 1645–1655 (1995).
- Peichel, C. L. et al. The genetic architecture of divergence between threespine stickleback species. *Nature* **414**, 901–905 (2001).
- Luttbeg, B., Townner, M. C., Wandesforde-Smith, A., Mangel, M. & Foster, S. A. State-dependent mate-assessment and mate-selection behavior in female threespine sticklebacks (*Gasterosteus aculeatus*, Gasterosteiformes: Gasterosteidae). *Ethology* **107**, 545–558 (2001).
- Raufaste, N. & Rousset, F. Are partial Mantel tests adequate? *Evolution* **55**, 1703–1705 (2001).
- Moore, A. Ø., Vamori, S. M. & Schluter, D. Using phylogenies to test macroevolutionary hypotheses of trait evolution in cranes (Gruinae). *Am. Nat.* **154**, 249–259 (1999).
- Hatfield, T. Genetic divergence in adaptive characters between sympatric species of stickleback. *Am. Nat.* **149**, 1009–1029 (1997).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank L. Bauers, V. Braithwaite, K. Faller, S. Foster, S. Gray, M. Ishikawa, P. Jacobsen, P. Katz, R. King, B. Kristjansson, M. Nemethy, H. Ogawa, W. Paulson, J. Poole, E. Sassman, S. Shell, S. Skulason and R. Snyder for help collecting fish and/or data. M. Blows, H. Rundle and A. Hendry provided useful comments on the manuscript and the University of Queensland hosted J.S.M. during writing. This project was supported by an NSF research grant, REU supplements and a Putnam grant (J.S.M.), the Howard Hughes Medical Institute (D.M.K.) and an NSERC grant (D.S.).

Authors' contributions Stickleback collection was conducted/supervised by J.S.M., S.M. and D.S.; mating trials by J.S.M., J.C. and L.J.; molecular work by B.K.B., L.D. and D.M.K.; and data analyses by J.S.M., D.S. and B.K.B.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.S.M. (mckinnoj@uww.edu).

Zinc transporter LIV1 controls epithelial-mesenchymal transition in zebrafish gastrula organizer

Susumu Yamashita^{1,2*}, Chiemi Miyagi^{2*}, Toshiyuki Fukada^{2,3}, Naofumi Kagara², Yong-Suk Che¹ & Toshio Hirano^{1,2,3}

¹Laboratory of Developmental Immunology, Graduate School of Frontier Biosciences and ²Department of Molecular Oncology, Graduate School of Medicine, Osaka University, Suita, Osaka 565-0871, Japan

³Laboratory for Cytokine Signaling, RIKEN Research Center for Allergy and Immunology, Yokohama, Kanagawa 230-0045, Japan

* These authors contributed equally to this work

Vertebrate gastrulation is a critical step in the establishment of body plan. During gastrulation, epithelial-mesenchymal transition (EMT) occurs¹. EMT is one of the central events of embryonic development, organ and tissue regeneration, and cancer metastasis^{1,2}. Signal transducers and activators of transcription (STATs) mediate biological actions such as cell proliferation, differentiation and survival in response to cytokines and growth factors, in a variety of biological processes^{3–6}. STATs are also important in EMT during gastrulation, organogenesis, wound healing and cancer progression^{7–9}. We previously showed that STAT3 is activated in the organizer during zebrafish gastrulation and its activity is essential for gastrulation movements. The requirement for STAT3 is cell-autonomous for the anterior migration of gastrula organizer cells, and non-cell-autonomous for the convergence of neighbouring cells¹⁰. The molecular mechanisms of STAT's action in EMT, however, are unknown. Here we identify LIV1, a breast-cancer-associated zinc transporter protein^{11–13}, as a downstream target of STAT3 that is essential and sufficient for STAT3's cell-autonomous role in the EMT of zebrafish gastrula organizer cells. Furthermore, we demonstrate that LIV1 is essential for the nuclear localization of zinc-finger protein Snail, a master regulator of EMT^{1,2,14,15}. These results establish a molecular link between STAT3, LIV1 and Snail in EMT.

We isolated a complementary DNA encoding zebrafish *LIV1* (*LZT-Dr3*; Supplementary Fig. 1a) by a subtraction screening. *LIV1* belongs to a subfamily of ZIP zinc transporters (Zrt-, Irt-like proteins), termed LZT (LIV-1 subfamily of ZIP zinc transporters)¹⁶. Its sequence contains eight transmembrane domains (TM) with a long extracellular amino terminus, a short extracellular carboxy terminus, a long variable region in the cytoplasmic loop between TM III and IV, an HNF motif in TM IV, a HEXPHE motif in TM V, and numerous histidine-rich repeats. All these motifs indicate

LIV1's relation to the LZT subfamily of ZIP zinc transporters¹⁶. The phylogenetic tree and alignment across the transmembrane domains IV and V of the LZT subfamily of ZIP zinc transporters revealed that the cDNA was the zebrafish counterpart of human *LIV1* (*LZT-Hs3*; Supplementary Fig. 1b). The zebrafish *LIV1* messenger RNA was expressed maternally, and zygotic transcripts increased at the shield stage and accumulated in the prechordal mesendoderm cells, in which STAT3 is activated¹⁰ (Fig. 1a–f). The expression of *LIV1* mRNA in the gastrula organizer was abolished in STAT3-depleted embryos, indicating that *LIV1* is a downstream target of STAT3 (Fig. 1g, h). Consistent with this, gp130-stimulation by G-CSF induced *LIV1* expression in the mouse proB cell line, Baf/B03 cells expressing a chimaeric receptor G-CSFR-gp130¹⁷ (Fig. 1i, G133), but not the Baf/B03 cells expressing the mutated chimaeric receptor defective in STAT3 activation (Fig. 1i, G133F3). Furthermore, the dominant negative form of STAT3 inhibited gp130-mediated *LIV1* expression in Baf/B03 cells (Fig. 1i, G133-STAT3F). Moreover, transfection of siRNA for human *STAT3* in the human prostate cell line, DU145 cells downregulated the expression of human *LIV1* (Fig. 1j). Together, these results showed that the *LIV1* gene is regulated by STAT3 not only in zebrafish organizer cells but also in some mammalian cells.

To address the early embryonic roles of *LIV1* during zebrafish gastrulation, we analysed embryos lacking *LIV1* activity using an antisense morpholino¹⁸ (*LIV1*-MO). Embryos receiving injections of *LIV1*-MO displayed a mispositioned head and shortened anterior–posterior axis at the end of gastrulation and later stages, whereas the injection of control *LIV1*D4-MO did not lead to any obvious phenotype (Fig. 2a–d). The axial hypoblast layer (notochord and somite) was formed in *LIV1*-MO-injected zebrafish embryos, but was shortened and thickened, suggesting that anterior movement of the axial mesendoderm was severely impaired, whereas the internalization, epiboly, and dorsal convergence move-

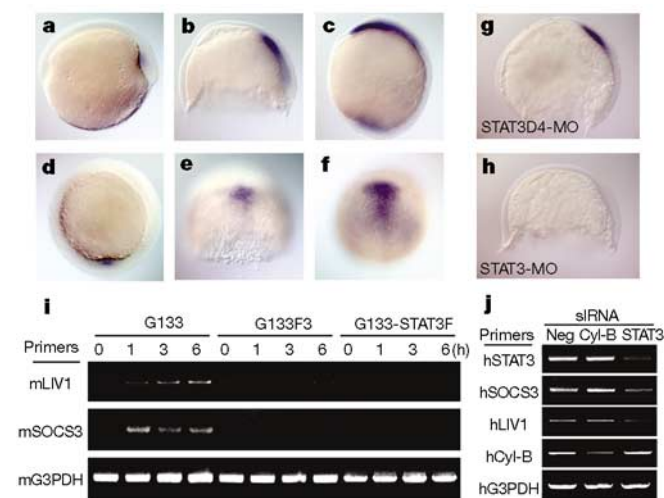


Figure 1 Zinc transporter *LIV1* is a target of STAT3 in zebrafish gastrula organizer cells. Zebrafish *LIV1* cDNA was isolated by a subtraction screening using normal embryos and STAT3-depleted embryos. **a–h**, *LIV1* mRNA at gastrula stage in normal embryo (**a–f**), STAT3D4 control morpholino-injected embryo (**g**), and STAT3 morpholino-injected embryo (**h**). **a–c**, **g**, **h**, Lateral views, with dorsal to the right; **d**, **f**, Animal-pole views, with dorsal to the bottom; **e**, dorsal views. **i**, Baf/B03 transfectant cells expressing the chimaeric receptor (G133), mutated chimaeric receptor defective in STAT3 activation (G133F3), or both the receptor and dominant negative STAT3 (G133-STAT3F) were cultured with G-CSF for the indicated times. Total RNA was isolated and subjected to RT–PCR with *LIV1*, *SOCS3* and *G3PDH* primers. **j**, DU145 cells were transfected with siRNA for either *STAT3*, negative control (Neg), or cyclophilin B (Cyl-B), followed by isolation of total RNA that was subjected to RT–PCR with primers *STAT3*, *SOCS3*, *LIV1*, cyclophilin B and *G3PDH*.

ments seemed normal. The expression of several marker genes highlighted the morphological abnormalities associated with the loss of LIV1 function (Fig. 2e–v). In LIV1-MO-injected embryos at the late gastrula stage, markers of the forebrain (*six3*), mid-hind-brain boundary (*pax2*), anterior axial mesoderm (*goosecoid*), posterior axial mesoderm (*no tail* and *axial*), paraxial mesoderm (*papc*), and endoderm (*axial*) were expressed, but the expression domain of these genes was vegetally mispositioned, and was shortened in the anterior–posterior axis and slightly expanded mediolaterally. Consistent with normal fate specification and regionalization in LIV1-deficient embryos at the late gastrula stages, normal expression of these mesendodermal genes was observed at the earlier stages of development (data not shown). These results indicated that LIV1 has essential roles during gastrulation by affecting cell movements, without significantly altering the early cell-fate specification of mesendodermal cells.

To clarify further how LIV1 affects cell movements during gastrulation, we performed cell-tracing experiments in zebrafish embryos using 4,5-dimethoxy-2-nitrobenzyl (DMNB)-caged fluorescein dextran¹⁹. In the control LIV1D4-MO-injected embryos, the marked cells were distributed in the dorsal axial hypoblast along the entire length of the anterior–posterior axis at the end of gastrulation (Fig. 3a, c). In contrast, in the LIV1-MO-

injected embryos, the anterior movement of the marked cells was perturbed, resulting in a shortened axial mesendoderm (Fig. 3b, c). However, marked cells in the lateral blastoderm margin, 90 degrees from the dorsal embryonic shield, moved dorsally and anteriorly, and extended along the anterior–posterior axis at the end of gastrulation regardless of LIV1-MO injection (Fig. 3d–f). These results clearly showed that LIV1 is essential for the active migration of organizer cells, but not for the dorsal convergence of non-axial mesodermal cells.

These observations suggested that *LIV1* is an essential target gene of STAT3 that is required for STAT3's cell-autonomous, but not non-cell-autonomous, roles. Indeed, when we cotransplanted prechordal mesendodermal cells from the embryonic shield of LIV1-MO-injected or LIV1D4-MO-injected control donor embryos into the embryonic shield of a normal or LIV1-depleted host embryo (Fig. 3g–k), prechordal mesendodermal cells from the control, but not the LIV1-depleted donors, migrated anteriorly to the animal pole in either normal (Fig. 3h) or LIV1-depleted host embryos (Fig. 3j). The reduced anterior migration of the LIV1-MO-injected cells was rescued by the co-injection of *LIV1* mRNA into the LIV1-depleted donor cells in either normal control (Fig. 3i) or LIV1-depleted (Fig. 3k) hosts. These results indicated that LIV1 acts cell-autonomously in the anterior migration of the prechordal mesendoderm.

We next sought to determine whether LIV1 is sufficient to induce the anterior migration of STAT3-depleted prechordal mesendodermal cells. STAT3-depleted cells, when transplanted into the embryonic shield, did not migrate anteriorly to the animal pole (Fig. 3l). This failure was rescued by the co-injection of *STAT3* mRNA (Fig. 3m), but not *STAT5* mRNA (Fig. 3n)¹⁰. Notably, *LIV1* mRNA rescued the cell-autonomous defect of the STAT3-depleted cells (Fig. 3o). This specific role of LIV1 was further indicated by the fact that *LIV1* mRNA could not rescue the failed non-cell-autonomous role of STAT3 in the gastrula organizer, that is, the induction of convergence (Supplementary Fig. 2a–e). All these data established that *LIV1* is essential and sufficient to carry out STAT3's cell-autonomous role, but not its non-cell-autonomous role, and suggested that LIV1 is a key molecule to regulate the mobility of organizer cells.

To clarify how LIV1 regulates the mobility of organizer cells, we next analysed the phenotypic change at the end of gastrulation and the migratory behaviour of organizer cells in LIV1-, STAT3- and Snail1-depleted embryos (Fig. 4a–h). During gastrulation, the organizer cells in normal embryos actively downregulate cell–cell adhesion systems, and leave their local neighbourhood to migrate individually, resulting in the body axis being fully extended anteriorly at the end of gastrulation (Fig. 4a, e). However, in LIV1-depleted embryos, the organizer cells could not downregulate their association, resulting in severe perturbation of the active migration of organizer cells (Fig. 4f), and a mispositioned head and shortened anterior–posterior axis (Fig. 4b). Similar defects in the migratory behaviour of organizer cells and body axis extension were also seen in STAT3- (Fig. 4c, g) and Snail1- (Fig. 4d, h) depleted embryos. Expression of *LIV1* mRNA and *Snail1* mRNA was normal in Snail1-MO-injected and LIV1-MO-injected embryos, respectively (Fig. 4i–p). Together, these results indicated that LIV1 and Snail1 are essential for EMT during gastrulation, and that the expression of *LIV1* and *Snail1* are independently regulated in gastrula organizer cells.

Consistent with this, the cell-autonomous defect in the anterior migration of Snail1-depleted organizer cells could not be rescued by the co-injection of *LIV1* mRNA (Fig. 3r), and vice versa (data not shown), suggesting that LIV1 might affect the activity of Snail. As shown in Fig. 4q, both *Snail* mRNA (lanes 1–4) and *LIV1* mRNA (lanes 6–8) repressed transcription from the Snail-responsive reporter plasmid in a dose-dependent manner. In addition, the co-injection of *LIV1* mRNA with a small amount of *Snail* mRNA

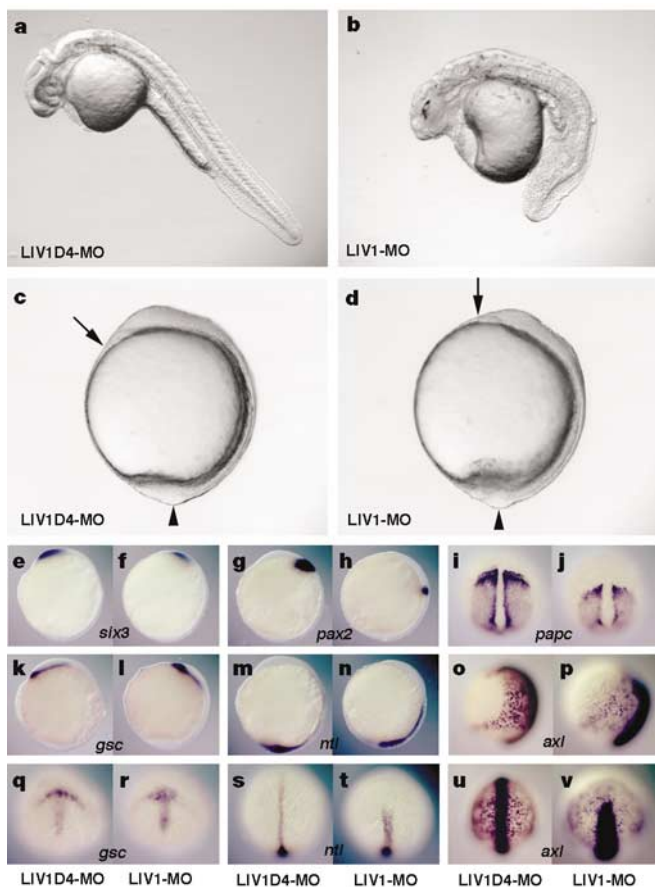


Figure 2 Disruption of gastrulation movement in LIV1 morphants. LIV1D4 control morpholino-injected, and LIV1 morpholino-injected embryos are shown. **a–d**, Phenotypes 1 day after fertilization (**a**, **b**, side views, anterior to the left) and the end of gastrulation (**c**, **d**, lateral views, dorsal to the right). Arrows in **c** and **d** mark the anterior limit of the hypoblast layer; arrowheads indicate the posterior limit of the hypoblast layer. **e–v**, Expression of *six3* (**e**, **f**), *pax2* (**g**, **h**), *papc* (**i**, **j**), *goosecoid* (**k**, **l**, **q**, **r**), *no tail/brachyury* (**m**, **n**, **s**, **t**), *axial/foxA2* (**o**, **p**, **u**, **v**). Panels **e–h** and **k–p** show lateral views, dorsal to the right, at the one-somite stage. Panels **i** and **j** and **s–v** show dorsal views at the one-somite stage. Panels **q** and **r** show animal-pole views at the one-somite stage.

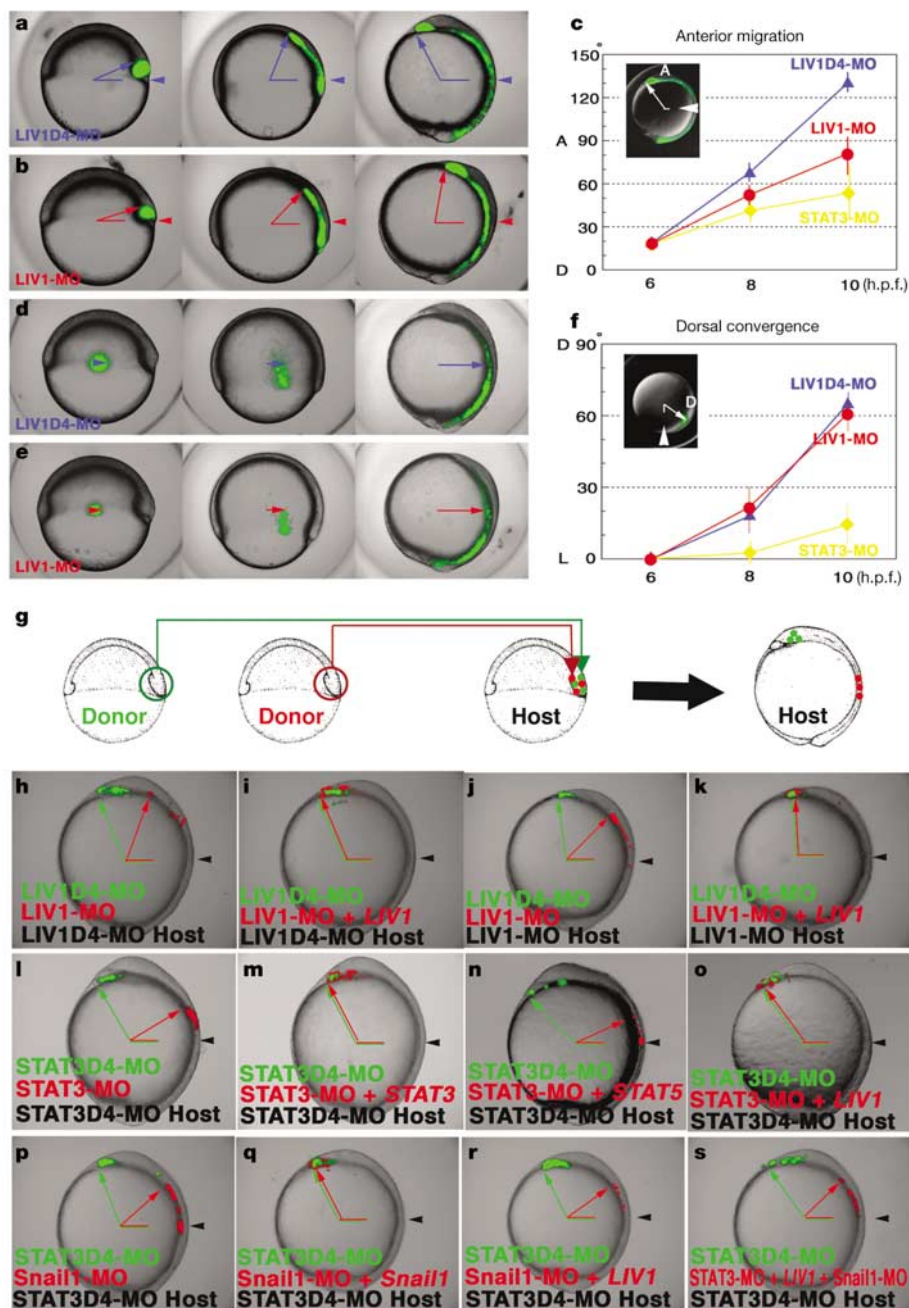


Figure 3 Dependence of cell-autonomous role of STAT3 in gastrulation movements on LIV1 activity. **a–f**, Requirement of LIV1 for anterior migration of gastrula organizer cells, but not for dorsal convergence of neighbouring cells. Axial mesendodermal cells in a control embryo (**a**) and a LIV1-depleted embryo (**b**). Lateral mesendodermal cells in a control embryo (**d**), and a LIV1-depleted embryo (**e**). Animal pole is up and dorsal to the right. Graphs compare anterior migration (**c**) and dorsal convergence (**f**) of labelled cell groups in control (LIV1D4-MO, blue), LIV1-depleted (LIV1-MO, red), and STAT3-depleted (STAT3-MO, yellow) embryos. Arrows indicate the leading edge of anterior migrating axial mesendodermal cells (**c**) and dorsal converging lateral mesendodermal cells (**f**); arrow-

heads indicate initial position at the shield stage. Each experiment was carried out ten times. The mean and standard deviation are shown. **g–s**, Prechordal mesendodermal cells (green or red) from morpholino- (upright text) and mRNA- (italic text) injected donors were transplanted into the shield region of a morpholino-injected host at the early gastrula stage. Panels **h–s** show the hosts at the end of gastrulation. Arrows indicate the leading edge of transplanted anterior migrating cells, and arrowheads indicate initial transplant position. Each experiment was carried out at least ten times, and one experiment with typical results is shown.

enhanced the repressor activity of Snail, in a dose-dependent manner (Fig. 4q, lanes 9–11). Importantly, the effect of LIV1 was completely abolished in Snail1-MO-injected zebrafish embryos (Fig. 4q, lanes 12–14). Furthermore, the cell-autonomous rescue by LIV1 of the anterior migration defect of STAT3-depleted organizer cells was sensitive to Snail-MO (Fig. 3o, s), clearly showing that LIV1 activity is dependent on Snail.

Because activity of the Snail transcriptional repressor is regulated

by its intracellular location and its nuclear accumulation is dependent on C terminus containing zinc-finger domains²⁰, we examined whether LIV1 regulates subcellular location of Snail by using Snail protein fused with green fluorescent protein (Snail-GFP). In control embryos, ectopically expressed Snail accumulated in the nuclei of organizer cells and non-organizer cells (Fig. 4r, u), as previously reported in mammalian cells²⁰. This nuclear accumulation was not observed when C-terminus-deleted Snail-GFP was used (Fig. 4s, v).

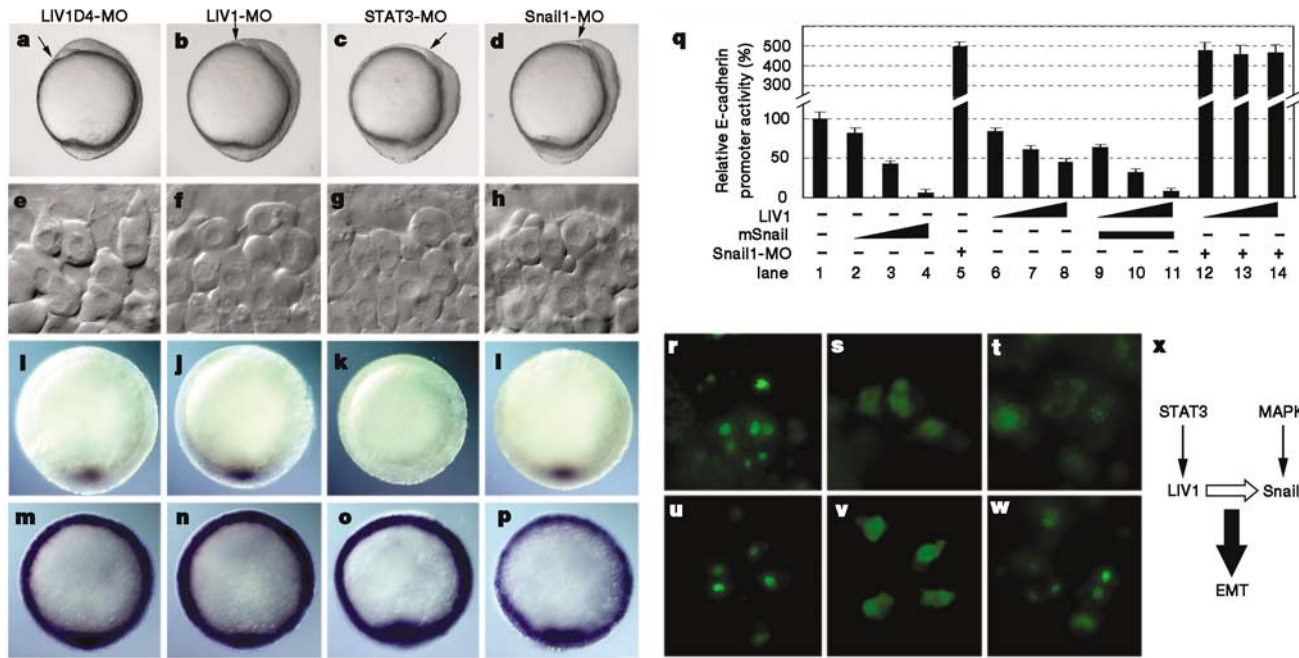


Figure 4 LIV1, STAT3 and Snail are essential for EMT of zebrafish gastrula organizer cells. **a–d**, Phenotypes of LIV1D4 control (**a**), LIV1- (**b**), STAT3- (**c**), and Snail1- (**d**) morpholino-injected embryos at the end of gastrulation (lateral views, dorsal to right). Arrow in **a–d** marks the anterior limit of the hypoblast layer. **e–h**, Morphology of organizer cells in the anteriormost limit of the hypoblast layer at mid-gastrula (8 h.p.f.) in control (**e**), LIV1-morphant (**f**), STAT3-morphant (**g**), and Snail1-morphant (**h**) embryos. **i–p**, Expression of *LIV1* (**i–l**) and *Snail1* (**m–p**) at early gastrula (6 h.p.f.) in control (**i**, **m**), LIV1-morphant (**j**, **n**), STAT3-morphant (**k**, **o**), and Snail1-morphant (**l**, **p**) embryos. Panels **i–p** show animal-pole views, with dorsal to the bottom. **q**, Snail-dependent LIV1 activity blocks *E-cadherin* promoter activity. One-cell-stage zebrafish embryos received co-injections of pGL3 vector containing the *Luc* gene under control of the pGL3-E-cadherin promoter plus control reporter vector (pRLtk) (lane 1). *Luc* activity was normalized to the *Rluc* activity. Co-injections were supplemented with zebrafish *LIV1* RNA (1, 10 or 100 pg per embryo, lanes 6–14), mouse *Snail* RNA (1, 10 or 100 pg per embryo, lanes 2–4, 9–11), and/or zebrafish

Snail1 morpholino (10 ng per embryo, lanes 5, 12–14). *Luc* activity was expressed as a percentage of the activity found in embryos receiving co-injections of control RNA and morpholino. The results show the average of three independent experiments, each performed in triplicate. The error bars indicate standard deviation.

r–w, LIV1-dependent nuclear translocation of Snail in zebrafish gastrula organizer cells. **r**, Snail-GFP in normal gastrula organizer cells. **s**, C-terminus-deleted Snail-GFP in normal gastrula organizer cells. **t**, Snail-GFP in LIV1-depleted gastrula organizer cells. **u**, Snail-GFP in normal non-organizer cells. **v**, C-terminus-deleted Snail-GFP in normal non-organizer cells. **w**, Snail-GFP in LIV1-depleted non-organizer cells. Frequency of cells showing nuclear localization of Snail ($n = 100$): 100% (**r**), 0% (**s**), 32% (**t**), 100% (**u**), 0% (**v**), and 97% (**w**). **x**, Model of the interactions between LIV1 and Snail in EMT of zebrafish gastrula organizer. STAT3 and MAPK independently induce the expression of *LIV1* and *Snail*, respectively. Zinc transporter LIV1 is essential for the nuclear translocation of zinc-finger protein Snail and affecting Snail transcriptional activity required for EMT.

In LIV1-depleted embryos, however, Snail protein showed extra-nuclear localization in organizer cells (Fig. 4t), but not in non-organizer cells (Fig. 4w), indicating that LIV1 is essential for nuclear localization of Snail protein in organizer cells. Together with the results of the phenotypic and gene expression analyses, these results provide evidence that LIV1 regulates the nuclear translocation of Snail and is involved in the induction of the EMT of organizer cells during zebrafish gastrulation.

Our results establish that LIV1, a downstream target of STAT3, plays a critical part in gastrula organizer cells undergoing EMT by affecting Snail activity. The *LIV1* gene, which encodes a zinc transporter protein, was originally identified as an oestrogen-regulated breast-cancer protein with possible involvement in metastatic spread^{11,12}. The *fear of intimacy* (*foi*) gene, involved in gonad morphogenesis in *Drosophila* has recently been found to encode a transmembrane protein closely related to human LIV1²¹. Previous work has indicated roles for the Snail/Slug family of repressors in EMT during cancer progression and gastrulation^{1,2,14,15}.

Although zebrafish *Snails* are expressed in both axial and non-axial mesendoderm^{22,23}, active anterior migration is observed only in the anterior axial mesendoderm²⁴, indicating the involvement of additional unknown mechanisms in active migration. The STAT3-dependent expression of *LIV1* specifically in the anterior axial mesendoderm and the requirement for LIV1 for normal EMT are consistent with the co-operative roles of Snail and LIV1 in the

invasive behaviour of organizer cells during gastrulation. A similar mechanism might contribute to cancer progression, because LIV1 and Snail are reported to be involved in the metastatic spread of breast-cancer cells^{12,25} and because STAT3 is constitutively activated in many tumours, including breast cancer²⁶. Finally, we propose a model for EMT in zebrafish gastrula organizer cells as shown in Fig. 4x (refs 27, 28). Our studies clearly establish a molecular link between STAT3, LIV1, and Snail during the EMT of gastrula organizer cells in zebrafish embryos. □

Methods

cDNA subtraction

cDNA subtraction was carried out by a polymerase chain reaction (PCR)-based method using a PCR-Select cDNA subtraction kit (Clontech), according to the manufacturer's protocol. To isolate genes induced by STAT3 in the gastrula organizer, poly (A)⁺ RNA was prepared from STAT3-MO-injected mid-gastrula-stage embryos and from STAT3D4-MO-injected mid-gastrula-stage embryos.

Morpholinos, mRNAs and plasmids

Morpholinos were obtained from Gene Tools. The sequences were as follows (the lower-case letters indicate mismatched sequences in the control morpholino): LIV1-MO, 5'-CGGAAACAGCGCGAGTGTCTTTTGT-3', 5'-ACCGTGTGCAAGAAACGTCATCAT-3'; LIV1D4-MO, 5'-CGGAAAGAGCGAGTGTATTTATGT-3', 5'-AcgGTCTGCAAAACAAAGTgATgAT-3'; STAT3-MO, 5'-CTCAAGGTTTCAGATAATCGTCT-3', 5'-GCCATGTTGACCCCTTAATGTGTCG-3'; STAT3D4-MO, 5'-CTCAGGATTCAGATAAAACGTgCT-3', 5'-GCCGTGTgACCCCTTAAAGTgACG-3'; Snail1-MO, 5'-GTCCACTCCAGTTACTTTTCAGGAT-3', 5'-CATGCTGAAGTCTGAAGTTGATC-3'. Morpholino oligonucleotides were injected into the yolk of one-cell-stage embryos.

Synthetic mRNA was injected into the yolk of embryos previously injected with morpholino. The following plasmids were linearized, and sense-strand-capped mRNA was synthesized: pCS2 + zebrafish LIV1 (*NotI*, SP6), pCS2 + zebrafish STAT3 (*NotI*, SP6)¹⁰, and pCS2 + mouse Snail (*NotI*, SP6)¹⁴. The cDNA sequence of full-length mouse Snail (1–264)²⁰ and C-terminus-deleted mouse Snail (1–151)²⁰ was amplified by high-fidelity PCR using specific oligonucleotides primers that created an *EcoRI* site at the 3' end. The resulting PCR product was cloned into a unique *EcoRI* site located upstream of the ATG of the GFP gene in the pCS2 expression plasmid.

Cell culture, RNAi and RT-PCR

Mouse proB cell lines Baf/B03 stably expressing human G-CSF-gp130 chimaeric receptors were maintained as described¹⁷. Cells were starved with factor-free medium for 6 h followed by stimulation with medium containing 100 ng ml⁻¹ of human G-CSF for the indicated time. Human prostate-cancer cell line DU145 was cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). The small interfering RNA (siRNA) for human STAT3 and siFactor were purchased from Dharmacon and B-Bridge, respectively. RNA interference (RNAi) screens were performed according to the manufacturers' protocols. Primers used for reverse transcription (RT)-PCR were as follows: mLIV1; 5'-AAAAATCCTAGAACTAGTCTAGGGAAAGGA-3' (sense), 5'-CCTTCAGCTCTCTCGAGAGTAGCGCTGGC-3' (antisense), mSOCS3; 5'-ATGGTCACCCACAGCAAGTTT-3' (sense), 5'-TTAAAGTGGAGCATCATACTG-3' (antisense), mG3PDH; 5'-TGAAGTCTGGTGTGAACGGATTGGC-3' (sense), 5'-CATGTAGGCCATGAGGTCCACCAC-3' (antisense), hSTAT3; 5'-GATCTGAATGG AAACAACCAAG-3' (sense), 5'-GCACACTTAATTTTAAGCTG-3' (antisense), hLIV1; 5'-GCAATGGCGAGGAAGTTATCT-3' (sense), 5'-CTATTGTCTCTAGAAAGTGAG-3' (antisense), hSOCS3; 5'-ATGGTCACCCACAGCAAGTTT-3' (sense), 5'-TTAAAGC GGGGCATCGTACTG-3' (antisense), hCyclophilin-B; 5'-ATGCTGCGCCTCTCCG AAGCG-3' (sense), 5'-GTACTCTCTGGCGATGGCAAA-3' (antisense), hG3PDH; 5'-TGAAGTCTGGAGTCAACGGAT-3' (sense), 5'-CATGTGGGCCATGAGGTCCAC-3' (antisense).

Cell-tracing experiments

Cell-tracing experiments were performed essentially as described previously¹⁰. 100 pl of 0.5% DMNB-caged fluorescein-dextran (molecular weight 10,000, Molecular Probes) was injected into the yolk cell of one-cell-stage embryos previously injected with 10 ng of the indicated morpholino. To uncage the dye, a beam of ultraviolet light ($\lambda < 360$ nm), generated using a DAPI filter set, was directed for 1 s at the dorsal or lateral blastoderm margin.

Transplantation experiments

Transplantation experiments were performed essentially as described previously¹⁰. Donor embryos were injected into the yolk cell of one-cell-stage embryos with 100 pl of 0.5% rhodamine-dextran (m_r 10,000, Molecular Probes) or with 100 pl of 0.5% fluorescein-dextran (m_r 10,000, Molecular Probes), and 10 ng of the indicated morpholino. A small population of deep cells (10–30 cells) from the embryonic shield of donor embryos was transplanted into the embryonic shield of host embryos at the same developmental stage.

Reporter assays

Reporter assays were performed essentially as described previously¹⁴. The reporter construct pGL3-E-cadh promoter (2.5 pg), which corresponds to the –178 to +92 human *E-cadherin* promoter fragment, was injected into one-cell-stage embryos with or without the synthetic RNA and/or the antisense morpholino. Firefly luciferase (Luc) and *Renilla reniformis* luciferase (Rluc) activities were measured using the Dual Luciferase Reporter Assay System (Promega) at the shield stage (6 hours post fertilization, h.p.f.), according to the manufacturer's protocol. Luc activity was always normalized to the Rluc activity.

Received 19 February; accepted 5 April 2004; doi:10.1038/nature02545.
Published online 5 May 2004.

- Thiery, J. P. Epithelial-mesenchymal transitions in tumour progression. *Nature Rev. Cancer* **2**, 442–454 (2002).
- Savagner, P. Leaving the neighborhood: molecular mechanisms involved during epithelial-mesenchymal transition. *Bioessays* **23**, 912–923 (2001).
- Darnell, J. E. Jr. STATs and gene regulation. *Science* **277**, 1630–1635 (1997).
- Bromberg, J. & Darnell, J. E. Jr. The role of STATs in transcriptional control and their impact on cellular function. *Oncogene* **19**, 2468–2473 (2000).
- Hirano, T., Ishihara, K. & Hibi, M. Roles of STAT3 in mediating the cell growth, differentiation and survival signals relayed through the IL-6 family of cytokine receptors. *Oncogene* **19**, 2548–2556 (2000).
- Kamimura, D. & Hirano, T. In *Signal Transducers and Activators of Transcription (STATs): Activation and Biology* (eds Sehgal, P. B., Levy, D. E. & Hirano, T.) 155–175 (Kluwer Academic, Dordrecht, 2003).
- Sano, S. *et al.* Keratinocyte-specific ablation of Stat3 exhibits impaired skin remodeling, but does not affect skin morphogenesis. *EMBO J.* **18**, 4657–4668 (1999).
- Silver, D. L. & Montell, D. J. Paracrine signaling through the JAK/STAT pathway activates invasive behavior of ovarian epithelial cells in *Drosophila*. *Cell* **107**, 831–841 (2001).
- Yamashita, S. & Hirano, T. In *Signal Transducers and Activators of Transcription (STATs): Activation and Biology* (eds Sehgal, P. B., Levy, D. E. & Hirano, T.) 595–607 (Kluwer Academic, Dordrecht, 2003).
- Yamashita, S. *et al.* Stat3 controls cell movements during zebrafish gastrulation. *Dev. Cell* **2**, 363–375 (2002).
- Manning, D. L., Daly, R. J., Lord, P. G., Kelly, K. F. & Green, C. D. Effects of oestrogen on the expression of a 4.4 kb mRNA in the ZR-75-1 human breast cancer cell line. *Mol. Cell. Endocrinol.* **59**, 205–212 (1988).
- Manning, D. L. *et al.* Oestrogen-regulated genes in breast cancer: association of pLIV1 with lymph node involvement. *Eur. J. Cancer* **30A**, 675–678 (1994).
- Taylor, K. M., Morgan, H. E., Johnson, A., Hadley, L. J. & Nicholson, R. I. Structure-function analysis

- of LIV-1, the breast cancer-associated protein that belongs to a new subfamily of zinc transporters. *Biochem. J.* **375**, 51–59 (2003).
- Battle, E. *et al.* The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nature Cell Biol.* **2**, 84–89 (2000).
- Cano, A. *et al.* The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nature Cell Biol.* **2**, 76–83 (2000).
- Taylor, K. M. & Nicholson, R. I. The LZT proteins; the LIV-1 subfamily of zinc transporters. *Biochim. Biophys. Acta* **1611**, 16–30 (2003).
- Fukada, T. *et al.* Two signals are necessary for cell proliferation induced by a cytokine receptor gp130: involvement of STAT3 in anti-apoptosis. *Immunity* **5**, 449–460 (1996).
- Nasevicius, A. & Ekker, S. C. Effective targeted gene 'knockdown' in zebrafish. *Nature Genet.* **26**, 216–220 (2000).
- Kozlowski, D. J. & Weinberg, E. S. Photoactivatable (caged) fluorescein as a cell tracer for fate mapping in the zebrafish embryo. *Methods Mol. Biol.* **135**, 349–355 (2000).
- Dominguez, D. *et al.* Phosphorylation regulates the subcellular location and activity of the snail transcriptional repressor. *Mol. Cell. Biol.* **23**, 5078–5089 (2003).
- Van Doren, M. *et al.* Fear of intimacy encodes a novel transmembrane protein required for gonad morphogenesis in *Drosophila*. *Development* **130**, 2355–2364 (2003).
- Thisse, C., Thisse, B., Schilling, T. F. & Postlethwait, J. H. Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development* **119**, 1203–1215 (1993).
- Thisse, C., Thisse, B. & Postlethwait, J. H. Expression of snail2, a second member of the zebrafish snail family, in cephalic mesoderm and presumptive neural crest of wild-type and spadetail mutant embryos. *Dev. Biol.* **172**, 86–99 (1995).
- Solnica-Krezel, L., Stemple, D. L. & Driever, W. Transparent things: cell fates and cell movements during early embryogenesis of zebrafish. *Bioessays* **17**, 931–939 (1995).
- Blanco, M. J. *et al.* Correlation of Snail expression with histological grade and lymph node status in breast carcinomas. *Oncogene* **21**, 3241–3246 (2002).
- Bowman, T., Garcia, R., Turkson, J. & Jove, R. STATs in oncogenesis. *Oncogene* **19**, 2474–2488 (2000).
- Ciruna, B. & Rossant, J. FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. *Dev. Cell* **1**, 37–49 (2001).
- Peinado, H., Quintanilla, M. & Cano, A. Transforming growth factor beta-1 induces snail transcription factor in epithelial cell lines: mechanisms for epithelial mesenchymal transitions. *J. Biol. Chem.* **278**, 21113–21123 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank A. G. De Herreros for the pGL3-E-cadh promoter plasmid and pcDNA3-mm snail-HA plasmid, and many colleagues for providing reagents. We also thank R. Masuda and A. Kubota for secretarial assistance. This work was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology in Japan.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.H. (hirano@molonc.med.osaka-u.ac.jp). The zebrafish *LIV1* cDNA sequence has been deposited in the GenBank database under accession number AB126260.

Identification of the pollen determinant of S-RNase-mediated self-incompatibility

Paja Sijacic¹, Xi Wang^{2*}, Andrea L. Skirpan², Yan Wang³, Peter E. Dowd², Andrew G. McCubbin^{2*}, Shihshieh Huang⁴ & Teh-hui Kao^{1,2,3}

¹Integrative Biosciences Graduate Degree Program, ²Department of Biochemistry and Molecular Biology, ³Intercollege Graduate Program in Plant Physiology, The Pennsylvania State University, 403 Althouse Laboratory, University Park, Pennsylvania 16802, USA

⁴Mystic Research, Monsanto Company, 62 Maritime Drive, Mystic, Connecticut 06355, USA

* Present addresses: Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109, USA (X.W.); School of Biological Sciences and Center for Reproductive Biology, Washington State University, Pullman, Washington 99164, USA (A.G.M.)

Many flowering plants have adopted self-incompatibility mechanisms to prevent inbreeding and promote out-crosses¹. In the Solanaceae, Rosaceae and Scrophulariaceae, two separate genes at the highly polymorphic S-locus control self-incompatibility interactions: the *S-RNase* gene encodes the pistil determinant and the previously unidentified *S*-gene encodes the pollen deter-

minant²⁻⁴. S-RNases interact with pollen S-allele products to inhibit the growth of self-pollen tubes in the style. Pollen-expressed F-box genes showing allelic sequence polymorphism have recently been identified near to the *S-RNase* gene in members of the Rosaceae and Scrophulariaceae⁵⁻⁸; but until now have not been directly shown to encode the pollen determinant. Here we report the identification and characterization of *PiSLF*, an S-locus F-box gene of *Petunia inflata* (Solanaceae). We show that transformation of S_1S_1 , S_1S_2 and S_2S_3 plants with the S_2 -allele of *PiSLF* causes breakdown of their pollen function in self-incompatibility. This breakdown of pollen function is consistent with 'competitive interaction', in which pollen carrying two different pollen S-alleles fails to function in self-incompatibility^{1,9,10}. We conclude that *PiSLF* encodes the pollen self-incompatibility determinant.

To maintain self-incompatibility, genes encoding the pollen and pistil specificity determinants must be tightly linked at the S-locus. Otherwise, intergenic recombination would cause the breakdown of self-incompatibility by generating different pollen and pistil specificities within a single S-haplotype. We thus searched for the pollen S-gene of *P. inflata* on a previously constructed 328-kilobase (kb) bacterial artificial chromosome (BAC) contig that contains *S₂-RNase*^{11,12}. Sequence analysis of this contig revealed an open reading frame for 389 amino acids, located ~161 kb downstream of the *S₂-RNase* gene. The predicted protein contains an F-box motif at its amino terminus, and was named *PiSLF₂* (*P. inflata* S-locus F-box, with '2' indicating the allele number).

To ascertain whether *PiSLF* is expressed, an 888-base-pair (bp) fragment of the *PiSLF₂* coding region (without the F-box coding sequence) was used as a probe in RNA gel blot analysis (Fig. 1a). A ~1.4-kb transcript was detected in anthers, mature pollen and pollen tubes, but not in any of the other tissues examined. During pollen development, the transcript was first detected in stage-three anthers (containing microspores undergoing mitosis), peaked in stage-four anthers (containing immature bicellular pollen) and subsequently declined. Although this expression pattern is somewhat unexpected for a gene that may function during pollen-tube growth in the style, transcript abundance may not reflect protein levels and the pollen tubes obtained by *in vitro* germination may not reflect the situation *in vivo*. Genomic DNA blot analysis revealed an

S-haplotype-specific restriction-fragment length polymorphism (RFLP) (Fig. 1b).

We used reverse transcription polymerase chain reaction (RT-PCR) to isolate a ~1.3-kb complementary DNA fragment of *PiSLF₂* from total RNA of S_2 pollen; the cDNA sequence is identical to the sequence of the corresponding region of *PiSLF₂* in the 328-kb contig. To determine the degree of allelic sequence diversity of *PiSLF*, we used RT-PCR to obtain cDNAs for *PiSLF₁* and *PiSLF₃* from the total RNA isolated from S_1 and S_3 pollen, respectively. The cDNA obtained for each S-genotype was used as a probe in genomic DNA blot analysis, and generated the same RFLP pattern as shown in Fig. 1b (results not shown), confirming that these cDNAs correspond to alleles of *PiSLF*. The deduced amino acid sequences of the S_1 -, S_2 - and S_3 -alleles of *PiSLF* (Supplementary Fig. S1) show 10.3% to 11.6% allelic sequence diversity, which is lower than that of the S-RNase (18.9% to 26.4%)¹³, but higher than that of two other S-linked F-box proteins of *P. inflata* that we had previously identified (3.2% to 5.1% for A113 and 1.8% to 3.4% for A134)¹⁴.

To ascertain whether *PiSLF* encodes the pollen S-determinant, we took advantage of a well-documented cause of breakdown of pollen S-function in the Solanaceae, termed competitive interaction. Competitive interaction occurs in self-incompatible plants that carry two different S-haplotypes when one of the S-loci, or some part of it that contains the pollen S-allele, is duplicated, or if these plants become tetraploids^{1,9,10}. Among the pollen grains produced, those carrying two different pollen S-alleles (that is, heteroallelic pollen), but not those carrying two copies of the same pollen S-allele (that is, homoallelic pollen), fail to function in self-incompatibility. How different pollen S-alleles 'compete' to cause the breakdown of self-incompatibility function is not understood. We reasoned that if *PiSLF* is the pollen S-gene, then introducing its S_2 -allele into self-incompatible plants of S_1S_1 genotype should render the transgenic plants self-compatible. As illustrated in Fig. 2a, if an S_1S_1 plant carries one copy of the *PiSLF₂* transgene, half the pollen produced will carry the transgene. Upon self-pollination, pollen carrying the transgene will be compatible with the pistil because of competitive interaction between the endogenous S_1 -allele and the introduced S_2 -allele of *PiSLF*. In contrast, pollen not carrying the transgene will

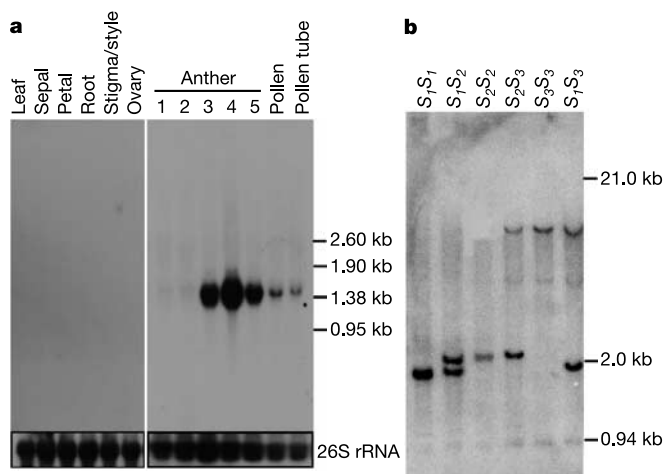


Figure 1 Characterization of *PiSLF*. **a**, RNA gel blot analysis of expression of *PiSLF*. Each lane contains 20 μ g of total RNA. The anther stages are defined by flower-bud size as previously described²¹. **b**, Genomic DNA blot analysis of six S-genotypes. Each lane contains 15 μ g of genomic DNA digested with *EcoRI*. The probe used for the upper blot in **a** and for the blot in **b** was an 888-bp fragment of *PiSLF₂* without the F-box motif region. DNA and RNA size markers are indicated.

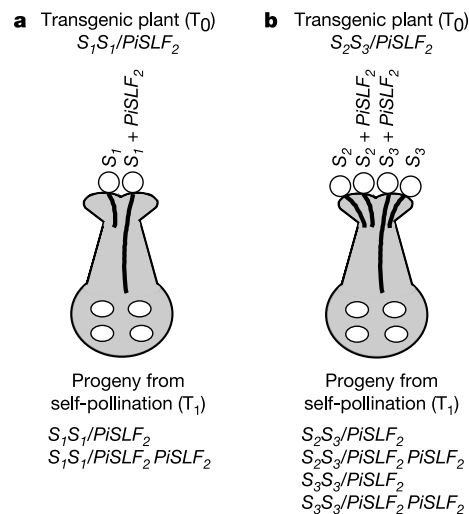


Figure 2 Schematic representation of transformation experiments to ascertain the function of *PiSLF*. **a**, Self-incompatibility behaviour of an S_1S_1 transgenic plant carrying a single copy of the *PiSLF₂* transgene. **b**, Self-incompatibility behaviour of an S_2S_3 transgenic plant carrying a single copy of the *PiSLF₂* transgene. The genotypes of pollen produced, the predicted S-genotypes of the progeny resulting from self-pollination, and inheritance of the transgene are indicated.

behave normally and be rejected by the pistil. Thus, all the progeny from self-pollination will carry the *PiSLF₂* transgene and be self-compatible.

A ~4.3-kb fragment of *PiSLF₂* (Supplementary Fig. S2) was introduced into *S₁S₁* plants via *Agrobacterium*-mediated transformation. Genomic DNA blot analysis showed that of the 25 *T₀* transgenic plants obtained, 20 (for example, the plant designated *S₁S₁/PiSLF₂-3*) carried two or more copies of the transgene, and five (designated *S₁S₁/PiSLF₂-1, -2, -4, -5, -6*) carried a single copy (Fig. 3a). The mature flowers of these five transgenic plants were self-pollinated and all pollinations resulted in large fruits, with an average seed number per fruit ranging from 115 to 170, comparable to those obtained from compatible pollination between wild-type plants. These five plants were further analysed to determine whether the breakdown of self-incompatibility was caused by competitive interaction.

First, we designed two pairs of PCR primers, one specific to *PiSLF₁* and the other to *PiSLF₂*, and used RT-PCR to examine the expression of both alleles of *PiSLF* in pollen. Products specific to *PiSLF₁* and *PiSLF₂* were obtained in all five plants (Fig. 3b), demonstrating that both endogenous *PiSLF₁* and the *PiSLF₂* transgene were expressed. This is consistent with the suggestion that competitive interaction is not due to silencing of pollen *S*-alleles present in heteroallelic pollen¹⁵.

Next, we examined whether it was the pollen or the pistil that failed to function in self-incompatibility. When a wild-type *S₁S₁*

plant was pollinated by pollen from *S₁S₁/PiSLF₂-1*, large fruits were obtained with an average seed number per fruit of 120. However, when *S₁S₁/PiSLF₂-1* was pollinated by pollen from the wild-type *S₁S₁* plant, no fruits were obtained. Similar results were found when *S₁S₁/PiSLF₂-2* was reciprocally crossed with the wild-type *S₁S₁* plant. Thus, the presence and expression of the *PiSLF₂* transgene caused the breakdown of the pollen function, but not of the pistil function, in self-incompatibility.

We raised 30 progeny (*T₁* plants) from self-pollination of *S₁S₁/PiSLF₂-1* and examined the inheritance of the *PiSLF₂* transgene. PCR analysis revealed that all 30 plants carried the *PiSLF₂* transgene (the results for eight *T₁* plants are shown in Fig. 3c). The presence of the transgene in these eight *T₁* plants was further confirmed by genomic blot analysis (Fig. 3d). The finding that all the progeny from self-pollination of *S₁S₁/PiSLF₂-1* inherited the *PiSLF₂* transgene is consistent with the breakdown of self-incompatibility resulting from competitive interaction between *PiSLF₁* and *PiSLF₂* (Fig. 2a). Moreover, all eight *T₁* plants were self-compatible and set large fruits with an average seed number per fruit ranging from 132 to 165. We also pollinated a wild-type *S₁S₁* plant with pollen from *S₁S₁/PiSLF₂-1* and -2, and found from PCR analysis that all 50 *T₁* plants of *S₁S₁/PiSLF₂-1* and all 49 *T₁* plants of *S₁S₁/PiSLF₂-2* carried the transgene (results not shown).

To further examine whether the *PiSLF₂* transgene specifically affects the self-incompatibility behaviour of heteroallelic pollen, we introduced the *PiSLF₂* transgene into *S₁S₂* and *S₂S₃* plants (the predicted results for *S₂S₃* are shown in Fig. 2b). Self-pollinations of three *S₁S₂* transgenic plants and three *S₂S₃* transgenic plants, all of which carried a single copy of the transgene, resulted in large fruits (with an average seed number per fruit ranging from 125 to 178). All of these plants rejected pollen from wild-type plants of the same *S*-genotype. The progeny from self-pollinations of *S₁S₂/PiSLF₂-1* and *S₂S₃/PiSLF₂-1* were analysed for the *S*-genotype and inheritance of the transgene. For *S₁S₂/PiSLF₂-1*, all 46 *T₁* plants were either *S₁S₁* (16) or *S₁S₂* (30); for *S₂S₃/PiSLF₂-1*, all 48 *T₁* plants were either *S₃S₃* (18) or *S₂S₃* (30). Figure 4 shows representative results of the progeny of *S₂S₃/PiSLF₂-1*. The absence of *S₂S₂* genotype in the progeny of these two transgenic plants suggests that *S₂* pollen was rejected, and hence the *PiSLF₂* transgene does not affect the self-incompatibility function of *S₂* pollen. Moreover, all of the 94 *T₁* plants examined carried the *PiSLF₂* transgene, suggesting that only *S₁* pollen carrying the transgene and *S₃* pollen carrying the transgene were compatible with *S₁S₂* and *S₂S₃* pistils, respectively.

From the results of the analyses of the *T₀* and *T₁* transgenic plants, we conclude that *PiSLF* encodes the pollen determinant of S-RNase-mediated self-incompatibility. The finding that an F-box protein

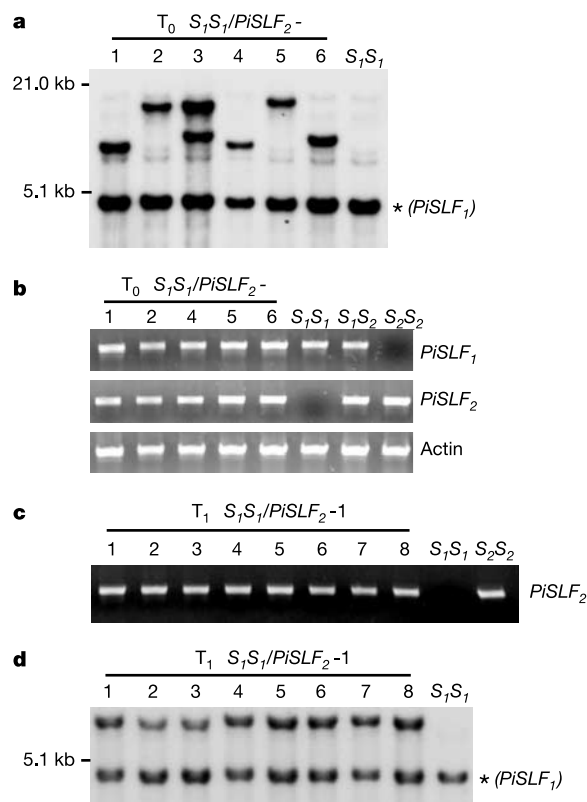


Figure 3 Analyses of *S₁S₁* transgenic plants carrying *PiSLF₂* transgene. **a**, Transgene copy number of six *T₀* plants. Genomic DNA was digested with *Xba*I and hybridized with the same *PiSLF₂* probe as that used in Fig. 1. **b**, RT-PCR analysis of endogenous *PiSLF₁* and the *PiSLF₂* transgene in the five *T₀* plants (*S₁S₁/PiSLF₂-1, -2, -4, -5, -6*) carrying a single copy of the transgene. **c**, PCR analysis of inheritance of the *PiSLF₂* transgene in eight *T₁* plants from self-pollination of *S₁S₁/PiSLF₂-1*. **d**, DNA blot analysis to confirm the results shown in **c**. In **a** and **d**, asterisks denote the endogenous *PiSLF₁* gene, and additional hybridizing fragment(s) correspond to the *PiSLF₂* transgene. Wild-type plants are included as controls.

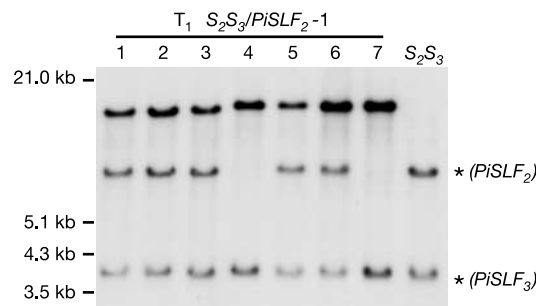


Figure 4 Analysis of progeny of an *S₂S₃* transgenic plant, *S₂S₃/PiSLF₂-1*, carrying a single copy of the *PiSLF₂* transgene. Genomic DNA was digested with *Xba*I and hybridized with the same *PiSLF₂* probe used in Fig. 1. Asterisks denote the endogenous *PiSLF₂* and *PiSLF₃* genes indicated in parentheses. The additional hybridizing fragment in each lane corresponds to the *PiSLF₂* transgene. A wild-type plant of *S₂S₃* genotype is included as a control.

determines pollen specificity in self-incompatibility interactions has important implications for how S-RNases specifically inhibit the growth of self-pollen tubes. It has been shown that RNase activity is required for the function of S-RNases¹⁶, and that both self and non-self S-RNases are taken up by pollen tubes growing in a pistil¹⁷. Therefore, it seems that only self S-RNase is able to function inside a pollen tube, degrading RNA substrate(s) to inhibit pollen-tube growth. This raises the question as to how the toxicity of non-self S-RNases is neutralized.

Recently, it has been shown that an S-locus F-box protein of *Antirrhinum* (Scrophulariaceae), AhSLF-S₂, is a likely component of the SCF complex (composed of Skp1, Cullin1, Rbx1 and an F-box protein)¹⁸, which is involved in ubiquitin-mediated protein degradation by the 26S proteasome¹⁹. Interestingly, AhSLF-S₂ interacts with both self and non-self S-RNases, but appears to mediate degradation of only non-self S-RNases¹⁸. Because there are multiple S-linked F-box genes in *Antirrhinum*⁸ and *P. inflata*¹⁴, whether AhSLF is an orthologue of PiSLF remains to be determined. If PiSLF is a component of the SCF complex, its interaction with a domain common to all S-RNases might lead to ubiquitination and degradation, but this interaction would be prevented by an alternate specific interaction between PiSLF and its cognate self S-RNase. □

Methods

DNA and RNA gel blot analyses

Isolation of genomic DNA from young leaves and genomic DNA blot analysis were carried out as previously described¹⁴, as was isolation of total RNA and RNA gel blot analysis²⁰.

Construction of transgene and *Agrobacterium*-mediated transformation

An ~11-kb DNA fragment containing PiSLF₂ was released from BAC clone 145J16 (ref. 12) by BamHI digestion and cloned into the BamHI site of pBluescript KS + (Stratagene) to yield pBS-11K. pBS-11K was digested with SalI, made blunt-ended by Klenow enzyme, and digested with XbaI to release a ~4.3-kb fragment containing PiSLF₂. A Ti plasmid, pBI101 (Clontech), was digested with SacI, made blunt-ended by T4 DNA polymerase, and digested with XbaI. The ~4.3-kb DNA fragment containing PiSLF₂ was ligated into the XbaI site and the blunt-end site of pBI101 (see Supplementary Fig. S2). The recombinant Ti plasmid was electroporated into *Agrobacterium tumefaciens* strain LBA4404 using a Cell-Porator (Life Technologies). The transformation of *P. inflata* leaf strips with *Agrobacterium* was carried out as described previously².

PCR and RT-PCR

To clone PiSLF₁, PiSLF₂ and PiSLF₃ by RT-PCR, cDNA was separately synthesized from 5 µg of total RNA isolated from S₁, S₂ and S₃ pollen using 200 units of SuperScript II RNaseH⁻ reverse transcriptase (Invitrogen). PCR was carried out on the cDNA of each S-genotype in a standard reaction buffer containing 0.25 µM each of forward primer (5'-GATCCAGTTGAATGAGATGG-3') and reverse primer, oligo (dT)₁₇, and two units of Taq DNA polymerase (Fisher Scientific). The reaction mixture was denatured at 94 °C for 1 min, then subjected to 30 cycles, with each cycle consisting of denaturation at 94 °C for 45 s, annealing at 45 °C for 45 s and extension at 72 °C for 1 min. After the last cycle, the sample was kept at 72 °C for an additional 10 min. PCR products were gel-purified and ligated into pGEM-T Easy vector (Promega).

To assess the presence of the PiSLF₂ transgene in T₀ and T₁ transgenic plants, PCR was carried out on 5 ng of genomic DNA isolated from each plant as described above, except that the forward primer was 5'-TATTAACATTCCAGTGAAATCTCTAT-3', the reverse primer was 5'-GTATCAGGCAATTTCATATCATGAAC-3' and the annealing temperature was 58 °C. To examine the expression of PiSLF₁ and PiSLF₂ in T₀ transgenic plants, cDNA was synthesized as described above. PCR conditions were the same as those described above except that for PiSLF₁, the annealing temperature was 63 °C and the primers used were: 5'-TGTTAACATTTCCTGTTAAATCTCTCT-3' (forward primer) and 5'-ATCTTTT ACTGGATCAATGAGCTTATTGG-3' (reverse primer). To use actin as the control in RT-PCR, two primers, 5'-GGCATCACACTTTCTACAATGAGC-3' (forward) and 5'-GATATCCACATCACATTTCATGAT-3' (reverse), were designed on the basis of the sequence of pollen-expressed *Arabidopsis* actin 12 (At3g46520).

Received 17 December 2003; accepted 30 March 2004; doi:10.1038/nature02523.

- de Nettancourt, D. *Incompatibility and Incongruity in Wild and Cultivated Plants* (Springer, Berlin, 2001).
- Lee, H.-S., Huang, S. & Kao, T.-h. S proteins control rejection of incompatible pollen in *Petunia inflata*. *Nature* **367**, 560–563 (1994).
- Murfelt, J., Atherton, T. L., Mou, B., Gasser, C. S. & McClure, B. A. S-RNase expressed in transgenic *Nicotiana* causes S-allele-specific pollen rejection. *Nature* **367**, 563–566 (1994).
- Kao, T.-h. & Tsukamoto, T. The molecular and genetic bases of S-RNase-based self-incompatibility. *Plant Cell* advance online publication 9 March 2004 (doi:10.1105/tpc.016154).
- Lai, Z. *et al.* An F-box gene linked to the self-incompatibility (S) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. *Plant Mol. Biol.* **50**, 29–42 (2002).

- Entani, T. *et al.* Comparative analysis of the self-incompatibility (S-) locus region of *Prunus mume*: identification of a pollen-expressed F-box gene with allelic diversity. *Genes Cells* **8**, 203–213 (2003).
- Ushijima, K. *et al.* Structural and transcriptional analysis of the self-incompatibility locus of almond: identification of a pollen-expressed F-box gene with haplotype-specific polymorphism. *Plant Cell* **15**, 771–781 (2003).
- Zhou, J. *et al.* Structural and transcriptional analysis of S-locus F-box (SLF) genes in *Antirrhinum*. *Sex. Plant Reprod.* **16**, 165–177 (2003).
- Golz, J. F., Su, V., Clarke, A. E. & Newbigin, E. A molecular description of mutations affecting the pollen component of the *Nicotiana glauca* S locus. *Genetics* **152**, 1123–1135 (1999).
- Golz, J. F., Oh, H.-Y., Su, V., Kusaba, M. & Newbigin, E. Genetic analysis of *Nicotiana glauca* pollen-part mutants is consistent with the presence of an S-ribonuclease inhibitor at the S locus. *Proc. Natl Acad. Sci. USA* **98**, 15372–15376 (2001).
- McCubbin, A. G., Zuniga, C. & Kao, T.-h. Construction of a binary artificial chromosome library of *Petunia inflata* and the identification of large genomic fragments linked to the self-incompatibility (S-) locus. *Genome* **43**, 820–826 (2000).
- Wang, Y. *et al.* Chromosome walking in the *Petunia inflata* self-incompatibility (S-) locus and gene identification in an 881-kb contig containing S₂-RNase. *Plant Mol. Biol.* (in the press).
- Ai, Y. *et al.* Self-incompatibility in *Petunia inflata*: isolation and characterization of cDNAs encoding three S-allele-associated proteins. *Sex. Plant Reprod.* **3**, 130–138 (1990).
- Wang, Y., Wang, X., McCubbin, A. G. & Kao, T.-h. Genetic mapping and molecular characterization of the self-incompatibility (S-) locus in *Petunia inflata*. *Plant Mol. Biol.* **53**, 565–580 (2003).
- Luu, D.-T. *et al.* Rejection of S-heteroallelic pollen by a dual-specific S-RNase in *Solanum chacoense* predicts a multimeric self-incompatibility pollen component. *Genetics* **159**, 329–335 (2001).
- Huang, S., Lee, H.-S., Karunanandaa, B. & Kao, T.-h. Ribonuclease activity of *Petunia inflata* S proteins is essential for rejection of self-pollen. *Plant Cell* **6**, 1021–1028 (1994).
- Luu, D.-T., Qin, K., Morse, D. & Cappadocia, M. S-RNase uptake by compatible pollen tubes in gametophytic self-incompatibility. *Nature* **407**, 649–651 (2000).
- Qiao, H. *et al.* The F-box protein AhSLF-S₂ physically interacts with S-RNases that may be inhibited by the ubiquitin/26S proteasome pathway of protein degradation during compatible pollination in *Antirrhinum*. *Plant Cell* **16**, 582–595 (2004).
- Bai, C. *et al.* Skp1 connects cell cycle regulation to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* **86**, 263–274 (1996).
- Skirpan, A. L. *et al.* Isolation and characterization of kinase interacting protein 1, a pollen protein that interacts with the kinase domain of PRK1, a receptor-like kinase of petunia. *Plant Physiol.* **126**, 1480–1492 (2001).
- Mu, J.-H., Lee, H.-S. & Kao, T.-h. Characterization of a pollen-expressed receptor-like kinase gene of *Petunia inflata* and the activity of its encoded kinase. *Plant Cell* **6**, 709–721 (1994).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank A. H. Omeis for the greenhouse work, J. Wang for routine laboratory assistance and the Monsanto Sequencing Facility for sequencing the BAC clones. This work was supported by a Predoctoral Fellowship for Students with Disabilities from the National Institutes of Health (P.E.D.) and by grants from the National Science Foundation (T.-h.K.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.-h.K. (txk3@psu.edu).

Self-incompatibility triggers programmed cell death in *Papaver* pollen

Steven G. Thomas & Veronica E. Franklin-Tong

School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Sexual reproduction in many angiosperm plants involves self-incompatibility (SI), which is one of the most important mechanisms to prevent inbreeding. SI is genetically controlled by the S-locus, and involves highly specific interactions during pollination between pollen and the pistil on which it lands. This results in the rejection of incompatible ('self') pollen, whereas compatible ('non-self') pollen is allowed to fertilize the plant¹. In *Papaver rhoeas*, S-proteins encoded by the stigma component of the S-locus interact with incompatible pollen, triggering a Ca²⁺-dependent signalling network^{2–7}, resulting in the inhibition of pollen-tube growth. Programmed cell death (PCD) is a

Table 1 Alleviation of SI-induced DNA fragmentation and tube-growth inhibition by DEVD

Treatment	DNA fragmentation (%)		Mean pollen-tube length (μm)	
	Incompatible pollen	Compatible pollen	Incompatible pollen	Compatible pollen
Control (untreated)	3.33 $\pm$ 1.76	4.00 $\pm$ 1.15	536.3 $\pm$ 78.9	546.7 $\pm$ 80.2
SI	71.58 $\pm$ 2.24	2.83 $\pm$ 0.60	162.9 $\pm$ 5.9	520.7 $\pm$ 78.6
SI + DEVD	18.97 $\pm$ 5.53	2.00 $\pm$ 0.00	376.0 $\pm$ 40.2	490.9 $\pm$ 27.8
SI + YVAD	52.00 $\pm$ 5.29	6.38 $\pm$ 3.43	160.8 $\pm$ 5.6	532.5 $\pm$ 73.6

SI, SI induction; SI + DEVD, DEVD treatment before SI induction; SI + YVAD, YVAD treatment before SI induction. Results are means $\pm$ s.e.m.; $n = 3$ in all cases.

mechanism used by many organisms to destroy unwanted cells in a precisely regulated manner^{8–10}. Here we show that PCD is triggered by SI in an S-specific manner in incompatible pollen. This provides a demonstration of a SI system using PCD, revealing a novel mechanism to prevent self-fertilization. Furthermore, our data reveal that the response is biphasic; rapid inhibition of pollen-tube growth is followed by PCD, which is involved in a later 'decision-making' phase, making inhibition irreversible.

Programmed cell death (PCD) has several key diagnostic features. These include nuclear DNA fragmentation, leakage of cytochrome *c* from the mitochondria into the cytosol, and cleavage of poly(ADP-ribose) polymerase (PARP)^{8,11,12}. A key enzyme involved in apoptosis in animal cells is caspase-3, which is activated in an autoprocessing cascade¹³ and is instrumental in the cleavage of nuclear DNA and PARP^{14,15}. Although it is generally accepted that apoptosis does not occur in plant cells, PCD in plants shares some apoptotic features. The tetrapeptide DEVD is an inhibitor of caspase-3 and has been used to provide evidence for a caspase-like activity in apoptotic animal cells and in plant cells undergoing PCD^{16–18}. We have used these markers and tools to investigate whether PCD is triggered in incompatible pollen as a consequence of the SI response.

We have demonstrated previously that DNA fragmentation is stimulated during SI in an S-specific manner⁷. Although our results indicated that PCD might be triggered in incompatible pollen, they were not conclusive. Here we present data demonstrating that DNA fragmentation was inhibited by pretreatment with the caspase-3 inhibitor I peptide, Ac-DEVD-CHO (DEVD). In incompatible pollen tubes that had been SI induced, a significantly increased incidence of DNA fragmentation, 21.5-fold that of untreated pollen, was observed (Table 1, $P < 0.001$). The same treatment of compatible pollen tubes revealed a highly significant difference in the DNA fragmentation compared with incompatible pollen tubes ($P < 0.001$), thereby demonstrating S-specificity. When DEVD was included before the SI treatment of incompatible pollen tubes, DNA fragmentation was significantly decreased in comparison with that induced by SI alone (Table 1, $P < 0.001$), and the DNA fragmentation was not significantly different from that found in untreated pollen tubes ($P = 0.054$ (n.s.)). We also used the caspase-1 inhibitor I peptide, Ac-YVAD-CHO (YVAD), as a potential negative control. Although less DNA fragmentation was detected in SI-induced incompatible pollen tubes pretreated with YVAD (Table 1), the effect of YVAD on the SI response was considerably less than that of DEVD ($P = 0.027$) and pollen tubes still exhibited 52.0% DNA fragmentation. Together, these data provide evidence that a DEVDase/caspase-like activity is triggered in an S-specific manner in incompatible pollen by the SI response, and that this mediates DNA fragmentation. Although the *Arabidopsis* genome sequence reveals no caspase homologues, reports of caspase-like activities in plant cells with sensitivity to animal caspase-3 inhibitors^{18,19} implicate their involvement in PCD.

We also examined the effects of these treatments on pollen-tube inhibition (Table 1, Fig. 1). Incompatible SI-induced pollen tubes were significantly shorter than compatible pollen tubes (Table 1; $P = 0.01$), which achieved the same length as the untreated incompatible pollen tubes (Table 1; $P = 0.896$ (n.s.)). Pretreatment with

DEVD markedly alleviated the pollen-tube inhibition stimulated by SI in incompatible pollen (Table 1; $P = 0.006$). However, pretreatment with YVAD had no significant effect on SI-induced pollen-tube inhibition (Table 1, $P = 0.802$ (n.s.)), and tube lengths were almost identical to those treated with S-proteins alone. These data therefore provide good evidence that a DEVDase/caspase-like activity is involved in SI-mediated pollen-tube inhibition.

As it is well established that SI arrests pollen-tube growth rapidly^{2,5}, the data also indicated that pollen-tube growth might be able to recommence, with the implication that DEVDase activity can be triggered after initial arrest. Monitoring growth rates confirmed this idea (Fig. 1). Pollen-tube growth arrested within 5 min of SI induction (Fig. 1a); pretreatments with DEVD and YVAD had no effect on growth rate (Fig. 1b). Pretreatment of incompatible pollen tubes with DEVD allowed growth after SI (Fig. 1c), whereas YVAD showed no alleviation of SI-induced inhibition. The DEVD–SI response was quite variable, but examination of individual pollen tubes (Fig. 1d) revealed that arrest occurred for between 15 and 45 min, after which pollen-tube growth reinitiated. Detailed examination of DEVD-treated pollen

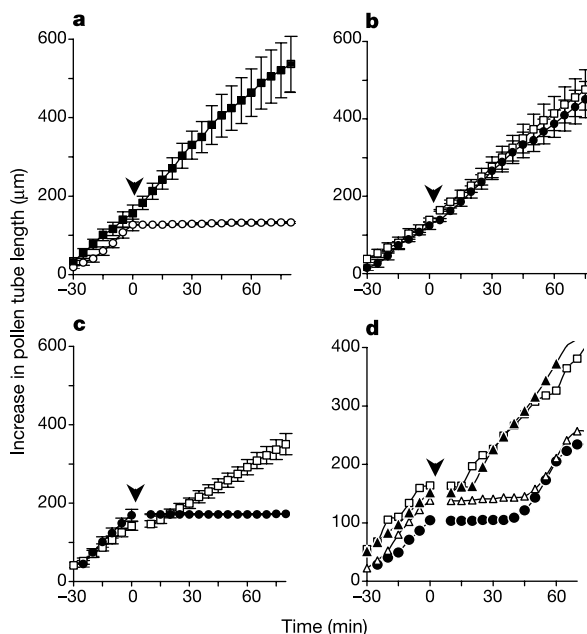


Figure 1 DEVD permits pollen-tube growth recovery after arrest by SI. **a**, SI induction (arrow) after pretreatment with GM (circles; $n = 7$) rapidly inhibited growth. Addition of GM instead of incompatible S proteins (squares; $n = 8$) at the arrow had no effect. Results are means $\pm$ s.e.m. **b**, Pretreatment of control pollen tubes with DEVD (squares; $n = 6$) or YVAD (circles; $n = 5$) before the addition of GM (arrow) had no effect on pollen-tube growth. Results are means $\pm$ s.e.m. **c**, Pollen tubes pretreated with DEVD followed by SI induction (squares; $n = 16$) grew after SI. Pretreatment with YVAD (circles; $n = 7$) before SI had no effect on inhibition. Arrow indicates SI induction. Results are means $\pm$ s.e.m. **d**, The growth of individual pollen tubes (represented by different symbols) responding to DEVD + SI was inhibited for between 15 and 45 min after SI; growth then recovered. The arrow indicates SI induction.

tubes between 0 and 10 min after SI revealed no growth during this period ($n = 19$; data not shown). Because YVAD, unlike DEVD, did not result in any significant pollen tube recovery, this supports the idea that it is the specific DEVD sequence that has this effect. Together, these data suggest not only that a DEVDase/caspase-like activity is activated by the SI response, but that it also has a crucial function in the inhibition of incompatible pollen-tube growth. Thus, there seems to be a biphasic SI response: rapid inhibition of pollen-tube growth, followed later by DEVDase activation, which presumably makes inhibition irreversible.

Cytochrome *c* leakage into the cytosol is a classic marker for PCD in many organisms, including plants^{20,21}. Here we present evidence that cytochrome *c* release from the mitochondria into the cytosol was stimulated by SI induction in incompatible pollen tubes (Fig. 2a). An increase in cytochrome *c* concentration was detected in cytosolic extracts from incompatible pollen 120 min after SI induction when compared with concentrations in compatible pollen tubes at the same time point (Fig. 2a); this demonstrates the S-specificity of this response. These data provide additional evidence that PCD is triggered in incompatible pollen during SI.

We investigated the timing of cytochrome *c* release into the cytosol, because this might give an indication of how early PCD is triggered in incompatible pollen (Fig. 2b). Small but detectable amounts of cytochrome *c* in incompatible pollen were observed as early as 10 min after SI induction. The concentration of cytosolic

cytochrome *c* increased further, and by 120 min after SI induction, amounts in the cytosol of incompatible pollen tubes were significantly higher than those in compatible pollen tubes at the same time point (Fig. 2b; $P = 0.001$). Controls consistently showed no sign of cytochrome *c* in the cytosol.

The cytosolic concentration of free Ca^{2+} ($[\text{Ca}^{2+}]_i$) acts as a second messenger in the SI response in *Papaver*². We investigated whether increases in $[\text{Ca}^{2+}]_i$ might also stimulate cytochrome *c* leakage. Mastoparan has been used previously to stimulate increases in $[\text{Ca}^{2+}]_i$ in *Papaver* pollen tubes²². Here we show that mastoparan stimulates the release of cytochrome *c* into the cytosol. A significant increase in cytosolic cytochrome *c*, from zero to $(9.92 \pm 1.4) \times 10^4$ counts mm^{-2} ($n = 3$, $P = 0.002$) was observed after 5 min of treatment with mastoparan. This indicates increases in $[\text{Ca}^{2+}]_i$ upstream of cytochrome *c* release, thereby providing a possible link with SI signalling.

PARP is a classic substrate for caspase-3 activity in animal cells and is often used as evidence for PCD. At least two PARP genes are known to be present in plants^{23,24}, so PARP could be a substrate for this activity. We examined SI-induced pollen protein extracts for a caspase-like cleavage activity *in vitro*, using bovine PARP as a substrate, and show that incompatible pollen extracts contain such an activity (Fig. 3a). A 24-kDa PARP cleavage fragment was consistently detected only in PARP samples incubated with incompatible SI-induced pollen extracts ($n = 8$). This, coupled with a

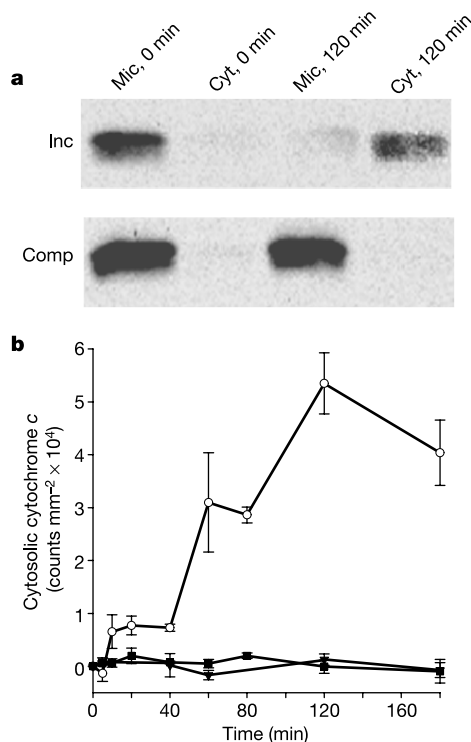


Figure 2 SI triggers the release of cytochrome *c* into the cytosol of incompatible pollen tubes. **a**, Microsomal (Mic) and cytosolic (Cyt) protein extracts from incompatible (Inc) and compatible (Comp) pollen that had undergone SI were examined for the presence of cytochrome *c* by using western blotting. Cytochrome *c* was present in microsomal, but not cytosolic, fractions from both compatible and incompatible pollen tubes at zero time. Increased amounts of cytochrome *c* were detected in the cytosolic fraction from incompatible pollen tubes 120 min after SI induction (lane 4, top panel), whereas compatible pollen tubes at the same time point showed no such increase (lane 4, bottom panel). **b**, After SI induction, the amount of cytosolic cytochrome *c* in incompatible (circles) pollen tubes increased over time. Compatible (triangles) and untreated (squares) pollen showed no change. Data are cytochrome *c* amounts (mean $\pm$ s.e.m.; $n = 3$) ascertained by quantification from western blots.

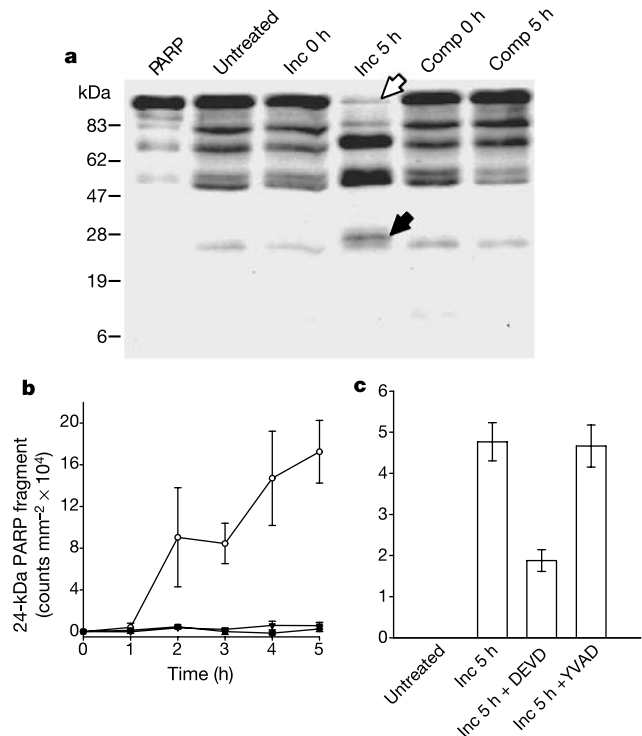


Figure 3 SI-stimulated incompatible pollen has an activity that can cleave bovine PARP. **a**, A PARP 24-kDa cleavage fragment (filled arrowhead) was detected in samples incubated with incompatible pollen extracts that had undergone SI for 5 h (Inc 5 h). A decrease in the full-length PARP protein (open arrow) was also evident. The PARP cleavage fragment was not detected in samples of PARP incubated without extracts (PARP) or with untreated, compatible (Comp) or incompatible (Inc) pollen extracts at zero time. **b**, PARP-cleavage activity in SI-induced incompatible (circles) pollen extracts increased in relation to the length of time after SI. Compatible (triangles) and untreated (squares) pollen extracts showed no PARP-cleavage activity. Results are amounts of 24-kDa PARP cleavage fragment (mean $\pm$ s.e.m.; $n = 3$) measured by quantification from western blots. **c**, Pretreatment of SI-induced pollen-tube extracts with DEVD (Inc 5 h + DEVD) decreased the amount of 24-kDa PARP cleavage fragment, whereas YVAD (Inc 5 h + YVAD) did not. Inc 5 h, SI-induced pollen-tube extracts.

corresponding decrease in the amount of uncleaved PARP (116 kDa) provides compelling evidence that a PARP-cleavage activity is triggered by SI. The S-specificity of this PARP-cleavage activity was demonstrated because neither compatible pollen that had been SI treated nor untreated pollen extracts generated a detectable 24-kDa PARP fragment (Fig. 3a). Investigation of the temporal activation of PARP-cleavage activity in incompatible pollen tubes by SI (Fig. 3b) revealed that the 24-kDa fragment was first detected 2 h after SI induction and it increased over time: 5 h after SI induction it was significantly higher than at zero time (Fig. 3b; $P = 0.005$). Because PARP is involved in nuclear DNA repair²⁵, the timing of detectable PARP-cleavage activity in incompatible pollen, which occurs before the detection of DNA fragmentation, is plausible.

We used mastoparan and La^{3+} to investigate whether Ca^{2+} might be implicated in stimulating PARP-cleavage activity in pollen tubes. Mastoparan stimulated a PARP-cleavage activity in pollen tubes, resulting in a highly significant increase in the abundance of the PARP 24-kDa cleavage product, $(2.88 \pm 0.366) \times 10^5$ counts mm^{-2} compared with zero in the untreated controls ($n = 4$, $P < 0.001$). We used La^{3+} , which blocks SI-stimulated Ca^{2+} influx in *Papaver* pollen tubes²⁶, to see whether this might provide further evidence for an involvement of Ca^{2+} . A comparison between the PARP-cleavage activity induced by SI in incompatible pollen tubes and the same treatment incubated with La^{3+} before SI induction resulted in a significant decrease ($70 \pm 7.7\%$, $n = 3$, $P = 0.002$) in the 24-kDa PARP fragment. Together these data indicate that increases in $[\text{Ca}^{2+}]_i$ in pollen can trigger PARP-cleavage activity, thereby providing a tangible link to Ca^{2+} signalling stimulated by SI.

We investigated whether the PARP-cleavage activity in incompatible pollen extracts might be due to a caspase-like activity by using DEVD as a competitive inhibitor and YVAD as a negative control (Fig. 3c). Pretreatment of SI-induced incompatible pollen extracts with DEVD before incubation with bovine PARP significantly decreased the amount of PARP 24-kDa cleavage product, a $60.38 \pm 5.66\%$ decrease ($n = 3$, $P < 0.001$) compared with that found in the same extracts in the absence of DEVD. The equivalent comparison with YVAD pretreatment gave a $0.64 \pm 14.23\%$ decrease ($n = 3$, $P = 0.966$). This clearly indicates the involvement of a DEVDase/caspase-like activity induced by SI in the generation of the 24-kDa PARP cleavage fragment.

In conclusion, we have examined the SI response in pollen of *P. rhoeas* for several diagnostic features of PCD. Together, our data provide compelling evidence that PCD is triggered by SI in an S-specific manner in incompatible pollen. The use of DEVD provides evidence for the involvement of a caspase-like activity in making the inhibition of incompatible pollen tube growth irreversible. The demonstration of PCD as a novel mechanism for the prevention of self-fertilization, together with evidence for a biphasic response, provides a significant advance in our understanding of SI and the regulation of pollen-tube growth. Because this SI system uses similar mechanisms to those used by the hypersensitive response in plant-pathogen interactions¹⁷, it provides evidence supporting previous speculative links between these recognition systems^{27,28}. □

Methods

SI induction and DNA fragmentation

Pollen was grown at 25 °C for 1 h on solid germination medium (GM)³ before SI induction, which was achieved by the addition of recombinant stigmatic S proteins S_{7e} and S_{8e} to S_{7s} (incompatible) and S_{7s} (compatible) pollen²⁹; pollen was incubated for a further 8 h. To test the effect of caspase inhibitors on SI-induced DNA fragmentation, pollen was grown for 1 h in the presence of 100 μM Ac-DEVD-CHO or Ac-YVAD-CHO (Biomol) before SI induction and incubation for 8 h. For 'untreated' controls, pollen was grown for 1 h, an aliquot of liquid GM (ref. 3) added, and pollen was incubated for a further 8 h. Times indicated are after SI induction. Pollen tubes were then fixed in 4% paraformaldehyde⁷ and labelled with a Deadend fluorimetric TUNEL kit (Promega).

Pollen tubes were scored for DNA fragmentation (50 tubes per treatment; $n = 3$) with a Nikon T300 fluorescence microscope. Capture and analysis of images was achieved with a Quips PathVision image analysis system (Applied Imaging International Ltd).

For the measurement of absolute pollen-tube lengths, pollen was treated as above but was incubated at 25 °C for 2 h, fixed in 4% paraformaldehyde and mounted in Tris-buffered saline (pH 7.6). Tube lengths were measured (60 tubes per treatment; $n = 3$) with the image analysis system described above and with IPlab software (Applied Imaging International Ltd). For growth rate analysis, increases in individual pollen-tube lengths were measured at 5 min intervals for different treatments (30 min before, and up to 80 min after) by using IPlab software for image capture and analysis.

Detection of cytochrome c in pollen-tube extracts

For SI treatments, pollen was grown in liquid GM (ref. 3) for 1 h at 25 °C and then SI was induced as described above. To investigate the role of Ca^{2+} in cytochrome c release, samples of pollen tubes were treated with mastoparan (25 μM , 5 min). After treatment, pollen tubes were incubated for various durations and centrifuged at 33g; mitochondrial extraction buffer³¹ was added to the pellet. Crude microsomal and cytosolic fractions were prepared as described in ref. 21. Equal concentrations of protein from both fractions were analysed by SDS-polyacrylamide-gel electrophoresis (SDS-PAGE) and western blotting³⁰. Blots were probed with a 1:2,000 dilution of a monoclonal antibody raised against a peptide from cytochrome c (clone 7H8.2C12; BD Biosciences), probed with an anti-mouse horseradish peroxidase secondary antibody (Sigma) and detected by enhanced chemiluminescence (ECL; Amersham Biosciences). Quantification of changes in the amount of cytochrome c in cellular fractions was established by determining the intensity of the cytochrome c signal on western blots by densitometry with a Fluor-S multi-imager and Quantity One software (Bio-Rad). Readings (counts mm^{-2}) were background-subtracted.

Measurement of PARP-cleavage activity in pollen-tube extracts

Pollen was grown in liquid GM (ref. 3) and SI was induced as described above. Pollen tubes were collected by centrifugation, and HEPES buffer (50 mM HEPES pH 7.4, 10 mM NaCl, 0.1% CHAPS, 10 mM dithiothreitol, 1 mM EDTA, 10% glycerol) was added. Samples were snap-frozen in liquid N_2 and pollen protein extracts were prepared by grinding, centrifuged (13,200 g for 20 min at 4 °C) and the supernatant was tested for PARP-cleavage activity. Pollen cytosolic protein (30 μg) was incubated with 250 ng of bovine PARP (Biomol) at 37 °C for 15 min. Final sample buffer ($5 \times$ SDS)³⁰ was added; samples were boiled for 10 min and analysed by SDS-PAGE and western blotting with a 1:2,000 dilution of monoclonal antibody raised against PARP (clone F1-23, Biomol), which detects the PARP 24-kDa fragment. Controls consisted of bovine PARP incubated with HEPES buffer, and untreated pollen extracts. Blots were probed and developed, and quantification of the 24-kDa PARP cleavage product was performed by densitometry as described above. To investigate the role of Ca^{2+} in the PARP-cleavage activity, samples of pollen tubes were treated with mastoparan (25 μM , 30 min); or with LaCl_3 (500 mM, 30 min) before SI induction. The effect of caspase inhibitors on PARP cleavage was investigated by the incubation of 30 μg of pollen cytosolic extracts with 100 μM Ac-DEVD-CHO or Ac-YVAD-CHO for 10 min before the addition of bovine PARP.

Received 27 November 2003; accepted 2 April 2004; doi:10.1038/nature02540.

- Franklin-Tong, V. E. & Franklin, F. C. H. Gametophytic self-incompatibility inhibits pollen tube growth using different mechanisms. *Trends Plant Sci.* **8**, 598–605 (2003).
- Franklin-Tong, V. E., Ride, J. P., Read, N. D., Trewavas, A. J. & Franklin, F. C. H. The self-incompatibility response in *Papaver rhoeas* is mediated by cytosolic free calcium. *Plant J.* **4**, 163–177 (1993).
- Rudd, J. J., Franklin, F. C. H., Lord, J. M. & Franklin-Tong, V. E. Increased phosphorylation of a 26-kD pollen protein is induced by the self-incompatibility response in *Papaver rhoeas*. *Plant Cell* **8**, 713–724 (1996).
- Rudd, J. J., Osman, K., Franklin, F. C. H. & Franklin-Tong, V. E. Activation of a putative MAP kinase in pollen is stimulated by the self-incompatibility (SI) response. *FEBS Lett.* **547**, 223–227 (2003).
- Geitmann, A., Snowman, B. N., Emons, A. M. C. & Franklin-Tong, V. E. Alterations in the actin cytoskeleton of pollen tubes are induced by the self-incompatibility reaction in *Papaver rhoeas*. *Plant Cell* **12**, 1239–1251 (2000).
- Snowman, B. N., Kovar, D. R., Shevchenko, G., Franklin-Tong, V. E. & Staiger, C. J. Signal-mediated depolymerization of actin in pollen during the self-incompatibility response. *Plant Cell* **14**, 2613–2626 (2002).
- Jordan, N. D., Franklin, F. C. H. & Franklin-Tong, V. E. Evidence for DNA fragmentation triggered in the self-incompatibility response in pollen of *Papaver rhoeas*. *Plant J.* **23**, 471–479 (2000).
- Raff, M. Cell suicide for beginners. *Nature* **396**, 119–122 (1998).
- Mittler, R. & Lam, E. Sacrifice in the face of foes: pathogen-induced programmed cell death in plants. *Trends Microbiol.* **4**, 10–15 (1996).
- Fukuda, H. Programmed cell death of tracheary elements as a paradigm in plants. *Plant Mol. Biol.* **44**, 245–253 (2000).
- Nagata, S., Nagase, H., Kawane, K., Mukae, N. & Fukuyama, H. Degradation of chromosomal DNA during apoptosis. *Cell Death Differ.* **10**, 108–116 (2003).
- Scovassi, A. I. & Poirier, G. G. Poly(ADP-ribosylation) and apoptosis. *Mol. Cell. Biochem.* **199**, 125–137 (1999).
- Salvesen, G. S. & Dixit, V. M. Caspase activation: the induced-proximity model. *Proc. Natl. Acad. Sci. USA* **96**, 10964–10967 (1999).
- Wolf, B. B., Schuler, M., Echeverri, F. & Green, D. R. Caspase-3 is the primary activator of apoptotic DNA fragmentation via DNA fragmentation factor-45/inhibitor of caspase-activated DNase inactivation. *J. Biol. Chem.* **274**, 30651–30656 (1999).
- Wolf, B. B. & Green, D. R. Suicidal tendencies: Apoptotic cell death by caspase family proteinases. *J. Biol. Chem.* **274**, 20049–20052 (1999).
- García-Calvo, M. et al. Inhibition of human caspases by peptide-based and macromolecular inhibitors. *J. Biol. Chem.* **273**, 32608–32613 (1998).

17. Michael, C., Lincoln, J. E., Bostock, R. M. & Gilchrist, D. G. Caspase inhibitors reduce symptom development and limit bacterial proliferation in susceptible plant tissues. *Physiol. Mol. Plant Pathol.* **59**, 213–221 (2001).
18. Danon, A., Rotari, V. I., Gordon, A., Mailhac, N. & Gallois, P. Ultraviolet-C overexposure induces programmed cell death in *Arabidopsis*, which is mediated by caspase-like activities and which can be suppressed by caspase inhibitors, p35 and defender against apoptotic death. *J. Biol. Chem.* **279**, 779–787 (2004).
19. Woltering, E. J., van der Bent, A. & Hoeberichts, F. A. Do plant caspases exist? *Plant Physiol.* **130**, 1764–1769 (2002).
20. Adrain, C. & Martin, S. J. The mitochondrial apoptosome: a killer unleashed by the cytochrome seas. *Trends Biochem. Sci.* **26**, 390–397 (2001).
21. Balk, J., Leaver, C. J. & McCabe, P. F. Translocation of cytochrome *c* from the mitochondria to the cytosol occurs during heat induced programmed cell death in cucumber plants. *FEBS Lett.* **463**, 151–154 (1999).
22. Franklin-Tong, V. E., Drobak, B. K., Allan, A. C., Watkins, P. A. C. & Trewavas, A. J. Growth of pollen tubes of *Papaver rhoeas* is regulated by a slow moving calcium wave propagated by inositol 1,4,5 trisphosphate. *Plant Cell* **8**, 1305–1321 (1996).
23. Babychuk, E. *et al.* Higher plants possess two structurally different poly(ADP-ribose) polymerases. *Plant J.* **15**, 635–645 (1998).
24. Lepiniec, L., Babychuk, E., Kushnir, S., Vanmontagu, M. & Inze, D. Characterization of an *Arabidopsis thaliana* cDNA homologue to animal poly(ADP-ribose) polymerase. *FEBS Lett.* **364**, 103–108 (1995).
25. Smith, S. The world according to PARP. *Trends Biochem. Sci.* **26**, 174–179 (2001).
26. Franklin-Tong, V. E., Holdaway-Clarke, T. L., Straatman, K. R., Kunkel, J. G. & Hepler, P. K. Involvement of extracellular calcium influx in the self-incompatibility response of *Papaver rhoeas*. *Plant J.* **29**, 333–345 (2002).
27. Dickinson, H. G. Simply a social disease? *Nature* **367**, 517–518 (1994).
28. Hodgkin, T., Lyon, G. D. & Dickinson, H. G. Recognition in flowering plants: a comparison of the *Brassica* self-incompatibility system and plant pathogen interactions. *New Phytol.* **100**, 557–569 (1988).
29. Kakeda, K. *et al.* Identification of residues in a hydrophilic loop of the *Papaver rhoeas* S protein that play a crucial role in recognition of incompatible pollen. *Plant Cell* **10**, 1723–1731 (1998).
30. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).

Acknowledgements We thank C. Franklin and B. de Graaf for a critical reading of the manuscript. This work was funded by the BBSRC.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to V.E.F.T. (v.e.franklin-tong@bham.ac.uk).

Rab5 is a signalling GTPase involved in actin remodelling by receptor tyrosine kinases

Letizia Lanzetti¹, Andrea Palamidessi^{1,2}, Liliana Areces^{1,2}, Giorgio Scita^{1,2} & Pier Paolo Di Fiore^{1,2,3}

¹Istituto Europeo di Oncologia, Via Ripamonti 435, 20141 Milan, Italy

²IFOM, Istituto FIRC di Oncologia Molecolare, Via Adamello 16, 20139 Milan, Italy

³Dipartimento di Medicina, Chirurgia ed Odontoiatria, Università degli Studi di Milano, 20122 Milan, Italy

Rab5 is a small GTPase involved in the control of intracellular trafficking, both at the level of receptor endocytosis and endosomal dynamics¹. The finding that Rab5 can be activated by receptor tyrosine kinases (RTK)² raised the question of whether it also participates in effector pathways emanating from these receptors. Here we show that Rab5 is indispensable for a form of RTK-induced actin remodelling, called circular ruffling. Three independent signals, originating from Rab5, phosphatidylinositol-3-OH kinase and Rac, respectively, are simultaneously required for the induction of circular ruffles. Rab5 signals to the actin cytoskeleton through RN-tre, a previously identified Rab5-specific GTPase-activating protein (GAP)³. Here we demonstrate that RN-tre has the dual function of Rab5-GAP and Rab5 effector. We also show that RN-tre is critical for

macropinocytosis, a process previously connected to the formation of circular ruffles^{4,5}. Finally, RN-tre interacts with both F-actin and actinin-4, an F-actin bundling protein. We propose that RN-tre establishes a three-pronged connection with Rab5, F-actin and actinin-4. This may aid crosslinking of actin fibres into actin networks at the plasma membrane. Thus, we have shown that Rab5 is a signalling GTPase and have elucidated the major molecular elements of its downstream pathway.

Activation of RTKs leads to the formation of actin-based structures known as membrane ruffles, through the activation of the small GTPases Ras and Rac^{5,6}. Ruffles are of at least two types: cell edge ruffles (lamellipodia) or dorsal surface ruffles (circular ruffles). Whereas lamellipodia are associated with cell motility/spreading, circular ruffles are involved in macropinocytosis and in three-dimensional migration^{4,5,7}. Active RTKs induce both lamellipodia and circular ruffles⁸. A linear cascade, RTK → Ras → Rac seems sufficient for lamellipodia formation^{9,10}, whereas the molecular events leading to circular ruffles are less clear. Active mutants of Ras or Rac do not induce circular ruffles^{9,10} (Supplementary Fig. 1), although an active form of Pak1, a Rac effector, does¹¹. In addition, Ras and Rac are required for the induction of circular ruffles because dominant negative mutants of these proteins inhibit their formation^{10,12}. Thus, concomitantly with the activation of Ras/Rac, additional RTK-triggered pathways may be required for circular ruffling.

Rab5 is a small GTPase required for the control of the endocytic route¹, and whose activity is regulated by guanine nucleotide exchange factors, such as Rabex5¹³, and GAPs, such as RN-tre³. Rab5 has been linked to RTK- and Ras-activated signalling^{2,14}, and in addition, an active Rab5 mutant induced actin remodelling¹⁵. Thus, Rab5 is a candidate regulator in the pathway leading to circular ruffles. Indeed, a dominant negative Rab5 mutant (Rab5^{S34N}) inhibited the induction of circular ruffles by platelet derived growth factor (PDGF), whereas cell edge ruffles were unaffected (Fig. 1a). Overexpression of the Rab5-specific GAP RN-tre³ also inhibited circular ruffles (Fig. 1a). Furthermore, Rab5 localized to PDGF-induced circular ruffles (Fig. 1b), similarly to Ras and Rac (Fig. 1c). Thus, the activity of Rab5 is required in a RTK-originated signalling pathway leading to circular ruffles. Notably, the effects of Rab5^{S34N} and RN-tre were not the consequence of inhibiting endocytosis, which can be hypothesized based

Table 1 Actin cytoskeleton remodelling activity of Rab5 in combination with various GTPases and effector proteins

Transfection or PDGF treatment	Membrane ruffles (%) [*]	Circular ruffles (%) [*]
Ras ^{V12}	+(89)†	–(0)
Rab5 ^{wt}	–(2)	–(0)
Rab5 ^{Q79L}	–(3)	–(0)
Ras ^{V12} + Rab5 ^{wt}	+(76)‡	+(23)‡
Ras ^{V12} + Rab5 ^{Q79L}	+(88)‡	+(14)‡
Ras ^{V12.S35}	–(2)	–(0)
Ras ^{V12.S35} + Rab5 ^{wt}	–(4)	–(0)
Ras ^{V12.C40}	+(93)‡	–(0)
Ras ^{V12.C40} + Rab5 ^{wt}	+(90)‡	+(30)‡
PI(3)K ^{CD2p110}	+(76)‡	–(0)
PI(3)K ^{CD2p110} + Rab5 ^{wt}	+(81)‡	+(22)‡
Rac ^{G61L}	+(88)‡	–(0)
Rac ^{G61L} + Rab5 ^{wt}	+(90)‡	–(0)
PDGF	+(91)‡	+(78)‡
PDGF + Rac ^{N17}	–(3)	–(0)
Ras ^{V12} + Rab5 ^{wt} + Rac ^{N17}	–(3)	–(0)
PI(3)K ^{CD2p110} + Rab5 ^{wt} + Rac ^{N17}	–(2)	–(0)

^{*}Actual images of selected experiments and further experimental controls are in Supplementary Figs 1 and 3. Unless otherwise specified (first column, PDGF), ruffles were scored under conditions of serum starvation. The phenotypes were detected (+) or not (–) in the transfected cells. Results are typical and represent at least three independent experiments, performed by transfection in MEFs.

†A quantification of the effects was performed in a single experiment, performed in triplicate in MEFs, by co-microinjection of the indicated plasmids. Phenotypes were scored by two observers. The numbers in parentheses (rounded up to the next higher integer) indicate the average number of cells displaying the phenotypes. At least 100 cells, in each triplicate, were counted.

‡Standard deviation was less than 20% of the mean. In the other samples, the number of events was too low for a meaningful statistical analysis.

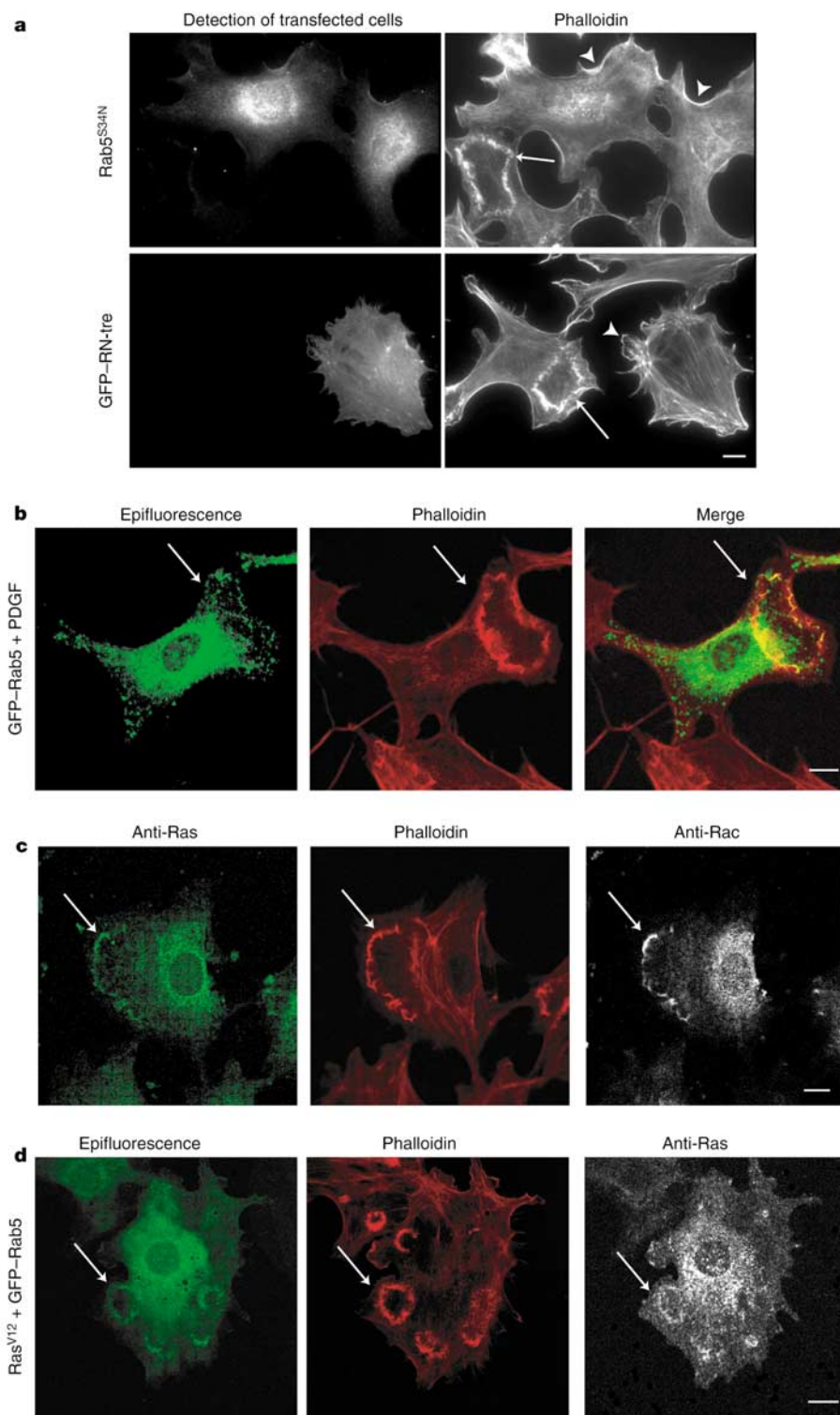


Figure 1 Rab5 and circular ruffles. **a**, MEFs, transfected with the constructs indicated on the left, were serum starved, PDGF-treated, and stained with phalloidin (right). Transfected cells (left) were detected with specific antibodies (top) or the epifluorescence of green fluorescence protein (GFP, bottom). Transfected cells displayed membrane ruffles (arrowheads), but no circular ruffles, which are detectable in approximately 80% of the untransfected cells (arrows). **b**, Confocal analysis of PDGF-treated/GFP-Rab5-expressing cells. GFP-Rab5 epifluorescence (green); TRITC-phalloidin (red); merge (yellow). **c**, Confocal analysis of PDGF-treated MEFs stained with anti-Ras antibody

(green), TRITC-phalloidin (red) and anti-Rac antibody (blue, pseudocoloured in black and white). **d**, Confocal analysis of serum-starved MEFs coexpressing Ras^{V12} and GFP-Rab5. Circular ruffles were detected in approximately 20% of the cotransfected cells, and in 0% of the mock-transfected cells. Rab5 epifluorescence (green); TRITC-phalloidin (red); Ras^{V12} using the anti-Ras antibody (blue, pseudocoloured in black and white). Circular ruffles are indicated with arrows in **b**, **c** and **d**. Scale bars, 10 μ m. Results are representative of at least three independent experiments.

on the role of Rab5 in this process. Indeed, a dominant negative mutant of the endocytic protein eps15 (eps15^{Δ95-295}), which acts by sequestering the endocytic adaptor AP2 and inhibits both ligand-induced and constitutive endocytosis¹⁶, did not cause any alteration in circular ruffling (Supplementary Fig. 2).

We dissected the pathway leading to circular ruffles. Overexpression of either wild-type Rab5 (Rab5^{wt}), or of its active mutant Rab5^{Q79L}, did not lead to actin remodelling (Table 1). However, coexpression of an active Ras (Ras^{V12}) with either Rab5^{wt} (Fig. 1d and Table 1) or Rab5^{Q79L} (Table 1), in serum-free medium, triggered the appearance of both cell edge and circular ruffles, whereas Ras^{V12} alone led to the formation of lamellipodia only (Table 1 and Supplementary Fig. 1). In Rab5^{wt}/Ras^{V12}-expressing cells, the formation of circular ruffles was more efficient than in Rab5^{Q79L}/Ras^{V12}-expressing cells (Table 1). Thus, in all subsequent experiments Rab5^{wt} was used.

Mutants of Ras^{V12} have been described that preferentially activate either one of the two major Ras-dependent pathways. Ras^{V12,S35} signals onto the Raf/MAPK pathway, whereas Ras^{V12,C40} signals to phosphatidylinositol-3-OH kinase (PI(3)K)¹⁷. Only Ras^{V12,C40} cooperated with Rab5 towards the induction of circular ruffles (Table 1). Moreover, coexpression of an active PI(3)K mutant (PI(3)K^{rCD2p110}) with Rab5 resulted in circular ruffle formation (Table 1). Thus, Rab5 is in a pathway parallel to the Ras/PI(3)K pathway and converges with it for the elicitation of circular ruffles.

Rac lies downstream of PI(3)K¹⁸. An active mutant of Rac (Rac^{Q61L}) induced lamellipodia formation, but showed no synergy with Rab5 towards circular ruffling (Table 1). However, a dominant negative mutant of Rac (Rac^{N17}) inhibited circular ruffles induced by PDGF, Ras^{V12}/Rab5 or PI(3)K^{rCD2p110}/Rab5 (Table 1). Thus, our data strongly support a model in which three independent signals,

originating from Rab5, PI(3)K and Rac, are simultaneously required for circular ruffling, whereas Rac signalling is sufficient for the induction of lamellipodia (see also Supplementary Fig. 3).

Clues as to the signals connecting Rab5 to actin remodelling have come from a structure-function analysis of RN-tre. RN-tre exerted a strong inhibition on circular ruffles (Figs 1a and 2a). A catalytic mutant (RN-tre^{R150}) displays no GAP activity on Rab5, and is inactive in bioassays in which wild-type RN-tre can suppress Rab5-dependent phenotypes³. Surprisingly, this mutant was still

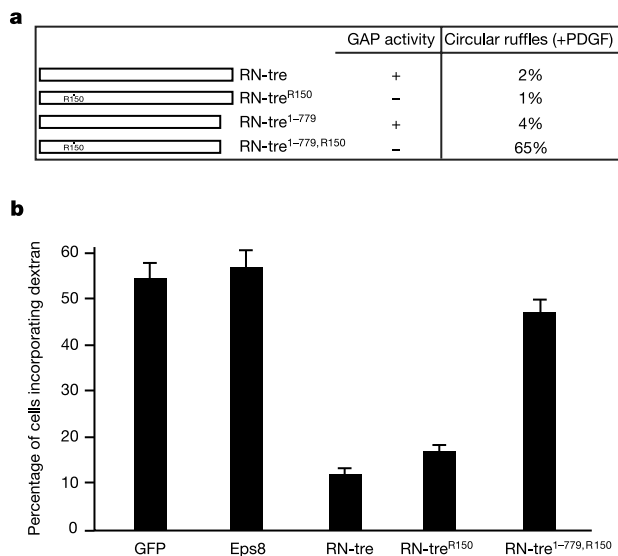


Figure 2 Biological effects of RN-tre and its mutants. **a**, A schematic representation of wild-type RN-tre and its mutants. The GAP activity was determined as described³. The effect on ruffling by the transfected GFP-tagged constructs (identified by epifluorescence) is reported as the per cent of cells (rounded up to the next higher integer) that showed circular ruffling when exposed to PDGF (see also Supplementary Fig. 4a). In cells transfected with a control vector and treated with PDGF, around 80% of the cells displayed circular ruffles. Data were obtained in a triplicate experiment, in which at least 100 cells per triplicate were scored. **b**, MEF cells transfected with the indicated plasmids (all encoding GFP-fusion proteins, GFP and GFP-Eps8 were used as negative controls) were serum starved and incubated with rhodamine-dextran and PDGF (10 ng ml⁻¹) for 30 min (see also Supplementary Fig. 4b). The percentage of transfected cells (identified by epifluorescence) that incorporated dextran is the average of triplicate samples $\pm$ standard deviation, and it is representative of three independent experiments.

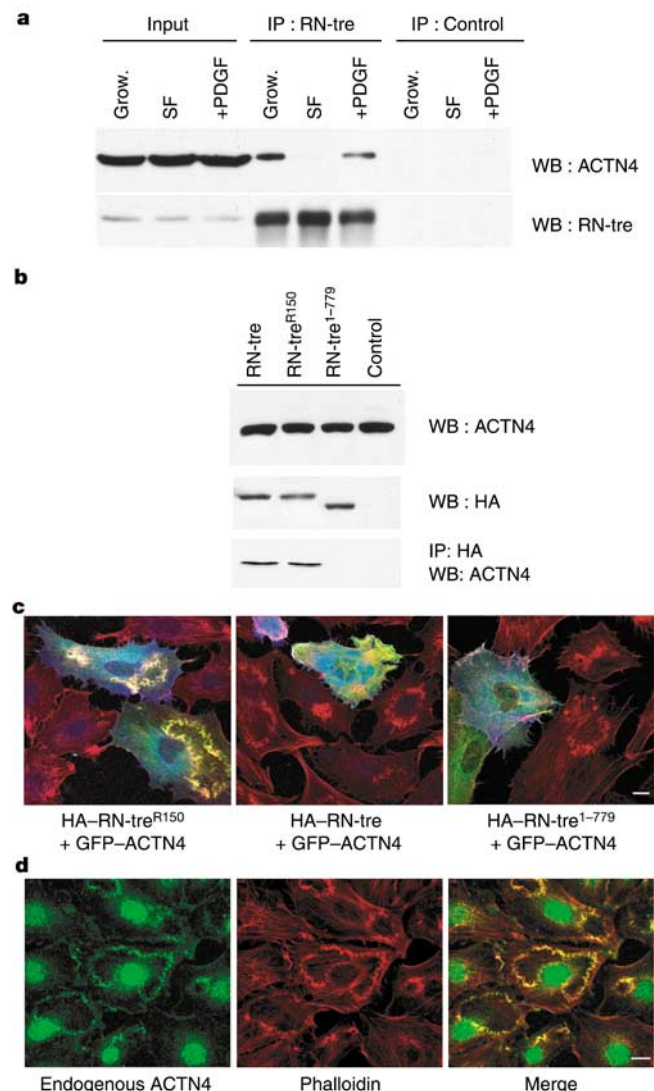


Figure 3 Interaction between RN-tre and ACTN4. **a**, MEF cells were lysed either at logarithmic growth (grow.), after serum starvation (SF) or after SF followed by PDGF treatment (20 ng ml⁻¹ for 15 min; +PDGF). Total lysates (50 μ g, input), anti-RN-tre (RN-tre) or anti-HA (control) immunoprecipitates (IP) (from 1.0 mg of lysates) were blotted (WB) with anti-ACTN4 (top) or anti-RN-tre (bottom) antibodies. **b**, 293T cells were transfected with the plasmids indicated (control, empty vector). All RN-tre constructs were HA-tagged. Cell lysates (100 μ g, top and middle panels) and anti-HA immunoprecipitates (from 1.0 mg of lysates, bottom panel) were western blotted with anti-HA or anti-ACTN4 antibodies. **c**, Confocal analysis of PDGF-treated MEF cells cotransfected with the plasmids indicated. GFP-ACTN4 epifluorescence (green); TRITC-phalloidin (red); anti-HA (blue). Percentage of cells displaying circular ruffles, quantification performed as in Table 1: RN-tre, 2%; RN-tre + ACTN4, 3%; RN-tre^{R150}, 2%; RN-tre^{R150} + ACTN4, 58%; RN-tre¹⁻⁷⁷⁹, 4%; RN-tre¹⁻⁷⁷⁹ + ACTN4, 5%; untransfected cells, 74% (all numbers are + PDGF). **d**, Confocal analysis of PDGF-treated MEF cells stained for endogenous ACTN4 (green), TRITC-phalloidin (red), merge (yellow). Scale bars, 10 μ m. Results are representative of at least three independent experiments.

able to suppress the formation of circular ruffles (Fig. 2a). A carboxy-terminal 49-amino-acid deletion mutant, RN-tre¹⁻⁷⁷⁹, suppressed circular ruffles and displayed GAP activity, as expected (Fig. 2a and Supplementary Fig. 4a). However, the double mutant RN-tre^{1-779,R150} had no GAP activity and could not inhibit ruffling (Fig. 2a and Supplementary Fig. 4a).

We attempted to provide physiological correlates for these findings. As circular ruffles have been implicated in macropinocytosis^{4,5}, we performed rhodamine-dextran (with a relative molecular mass of 70,000 (M_r 70K)) internalization assays. The ability of RN-tre mutants to inhibit circular ruffling perfectly correlated with their ability to inhibit macropinocytosis (Fig. 2b). Thus, two independent molecular functions of the RN-tre protein (that is, the GAP activity and another residing in the C terminus) can be linked to both the processes of circular ruffling and macropinocytosis (Supplementary Fig. 4b).

To understand the C-terminal function of RN-tre we identified cellular proteins interacting with this region by mass spectrometry. One of the interacting proteins was actinin-4 (ACTN4) (Supplementary Fig. 4c), a nonmuscle isoform of α -actinin¹⁹. ACTN4

is an actin-binding protein that crosslinks F-actin (see ref. 20 and references therein). Interestingly, ACTN4 has been preferentially implicated in actin bundling leading to circular ruffles, rather than to lamellipodia²¹. Endogenously expressed RN-tre and ACTN4 could be co-immunoprecipitated from cell lysates (Fig. 3a). This *in vivo* association was PDGF dependent, supporting the notion that it is functional in the execution of RTK-activated cellular programmes (Fig. 3a). Co-immunoprecipitation was strictly dependent on the presence of the last 49 amino acids of RN-tre (Fig. 3b).

Bacterially-expressed RN-tre fragments and ACTN4 did not associate *in vitro* (Supplementary Fig. 4d), suggesting that the interaction is indirect. We note, however, that the *in vivo* association is PDGF-dependent, a finding compatible with its dependence on signalling-induced post-translational modifications, most likely to be absent in the bacterially expressed proteins. Whatever the case, we succeeded in demonstrating the functional relevance *in vivo* of this interaction, in that the overexpression of ACTN4 rescued the inhibitory effect of the RN-tre^{R150} mutant on PDGF-induced circular ruffling (Fig. 3c). As expected, ACTN4 did not rescue the effects of RN-tre^{wt} or of the RN-tre¹⁻⁷⁷⁹ mutant (Fig. 3c), which still

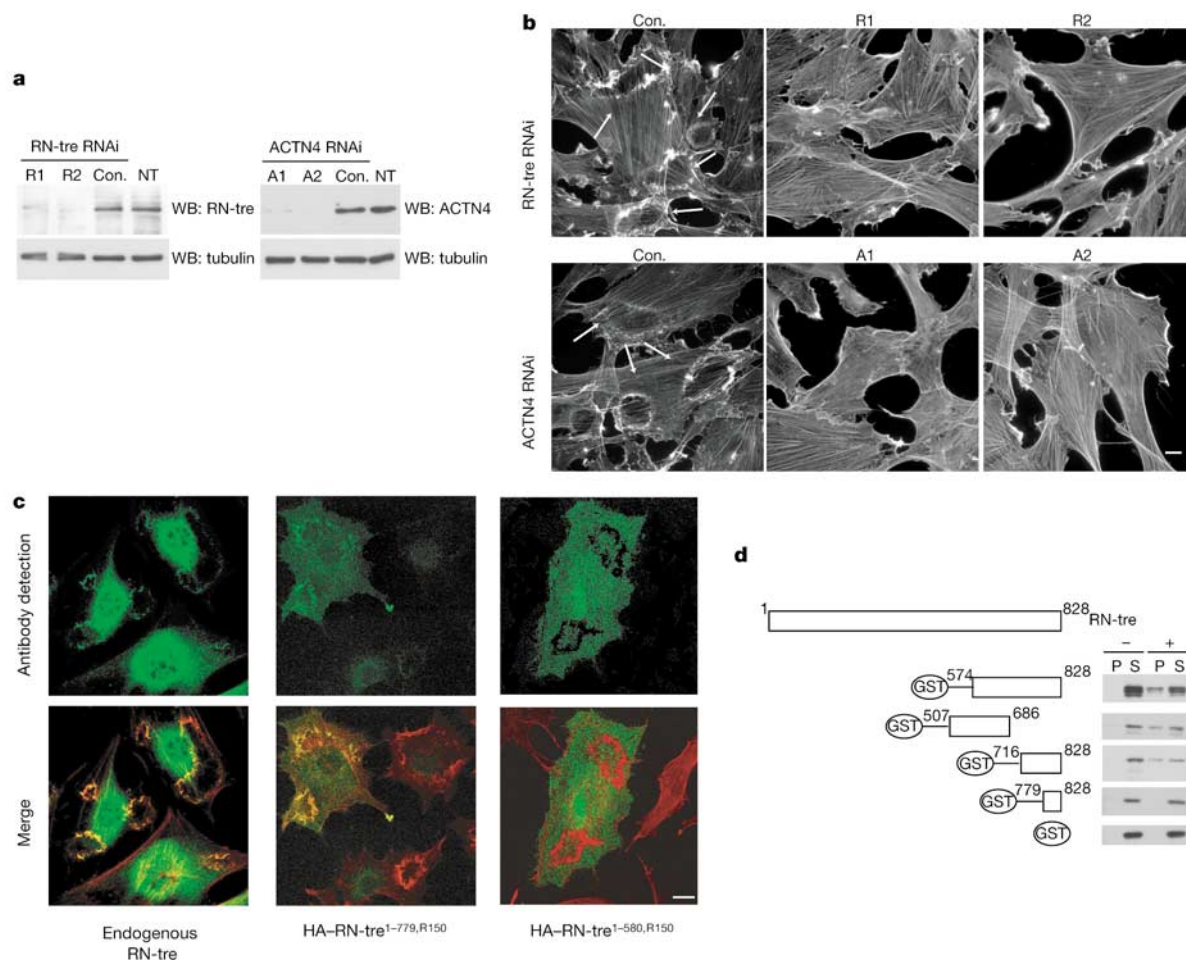


Figure 4 Role of RN-tre and ACTN4 in circular ruffling. **a**, Human fibroblasts were transfected with siRNA oligos for RN-tre (left panels, R1 and R2), ACTN4 (right panels, A1 and A2), control oligos (Con.) or untransfected (NT). Total lysates (20 μ g) were blotted (WB) with the indicated antibodies, and tubulin levels were used to verify equal loading. **b**, Cells were subjected to RNA interference as in **a** and serum starved for 4 h, treated with 10 ng ml⁻¹ of PDGF for 8 min and stained with phalloidin. Circular ruffles are indicated by arrows. Circular ruffles were scored in 42% of the untransfected and 40% of control-transfected cells and in 0% of the RN-tre- or ACTN4-silenced cells. At least 200 cells per triplicate experiment were counted. **c**, MEF cells were either mock transfected

(endogenous RN-tre) or transfected with the indicated RN-tre mutants, treated with PDGF and then stained with either anti-RN-tre (endogenous) or anti-HA (for the mutants) antibodies (antibody detection). Cells were co-stained with phalloidin (red) and the merged image is shown (yellow). Scale bars, 10 μ m. Results are representative of at least three independent experiments. **d**, The GST-RN-tre fragments used are shown, together with full-length RN-tre. In the co-sedimentation assays (right panels), the GST fusion proteins were incubated in the absence (–) or presence (+) of F-actin and then subjected to high-speed ultracentrifugation to separate pellets (P) and supernatants (S). Detection was performed with anti-GST antibodies.

displayed Rab5 GAP activity. Finally, endogenous ACTN4 localized to PDGF-induced circular ruffles (Fig. 3d).

These results strongly support the hypothesis that ACTN4 lies downstream of RN-tre and Rab5. A simple possibility to explain the connections among these three proteins is to postulate that RN-tre is an effector of Rab5. GAP proteins bind preferentially to the GTP-bound forms of small GTPases, thus fulfilling the requirement for effectors. In addition, there is precedent for GAPs behaving as GTPase effectors (for example ref. 22). A corollary of this hypothesis is that the ablation of RN-tre and of ACTN4 should result in the abrogation of circular ruffling. Indeed, RNA interference-mediated silencing of either RN-tre or of ACTN4 (Fig. 4a) resulted in a marked impairment of PDGF-induced circular ruffles, with cell edge ruffles unaltered (Fig. 4b). Similar results were obtained by microinjecting the cells with anti-RN-tre antibodies (Supplementary Fig. 4e).

We investigated whether RN-tre also interacts with actin. Endogenous RN-tre and the RN-tre^{1-779,R150} mutant localized to circular ruffles, whereas a further deletion mutant, RN-tre^{1-580,R150}, did not (Fig. 4c). This is compatible with the possibility that a region in RN-tre (amino acids 580–779), distinct from the ACTN4-interacting site (amino acids 779–828), is involved in binding to actin. Co-sedimentation assays of the recombinant C-terminal region of RN-tre^{574–828} with F-actin showed a direct interaction between the two proteins. Further mapping showed that at least two regions in RN-tre (amino acids 507–686 and 716–828) interacted directly with F-actin, whereas the ACTN4-interacting site (amino acids 779–828) did not (Fig. 4d). Together, our data strongly suggest that RN-tre has the potential to interact simultaneously with actin and ACTN4.

In this study, we demonstrate that Rab5 is a signalling protein involved in actin remodelling that is triggered by RTKs. Our results are compatible with a model in which Rab5 participates to the formation of circular ruffles (Rab5 localizes to these structures) through its effector RN-tre. The function of RN-tre may be that of mediating a three-pronged interaction with Rab5, ACTN4 and actin. This may help to localize ACTN4 to the proper plasma membrane location, where actin fibres are crosslinked into actin networks. Notably, our model accommodates for the need of multiple cycles of localization/crosslinking, in that the GAP activity of RN-tre may release Rab5 from the interaction with RN-tre, allowing for reloading of the GTPase with GTP and further cycles. The signalling-dependence of the interaction between RN-tre and ACTN4 also fulfils the requirement for reversibility of this interaction in our model. Finally, ACTN4 might represent a molecular intersection of pathways leading to circular ruffles, as it has been shown that phosphatidylinositol-3,4,5-trisphosphate (the catalytic product of PI(3)K activity) binds to α -actinin and modulates its activity, its distribution *in vivo*, and the flexibility of its actin-binding region^{23,24}. Notably, the phosphoinositide binding site is conserved in ACTN4 (refs 19, 25).

The physiological significance of circular ruffling and its molecular mechanisms has recently started to be elucidated. A dynamin–cortactin complex may be required in the formation of these structures (called ‘waves’)²⁶; a result that can be correlated to the *in vitro* property of dynamin to regulate the stimulation of Arp2/3-dependent actin polymerization by cortactin²⁷. In addition, it has been shown that within the WASP (Wiskott–Aldrich syndrome protein)-family Verprolin-homologous protein (WAVE) subfamily of proteins (which connects active Rac to the actin nucleating Arp2/3 complex), WAVE1 is required for circular ruffling, whereas WAVE2 regulates lamellipodia⁷. Interestingly, circular ruffling can be correlated to three-dimensional migration⁷. We note that macropinocytosis can also be regarded as an efficient way to recycle plasma membrane, which would be required in the process of cell migration. In this framework, our findings that Rab5 directs the process of circular ruffling raises the possibility that this GTPase, which has already been shown to delimit specialized membrane

domains in the endosome¹, may serve as a multi-purpose device that regulates specialized membrane-linked events in several cellular compartments. □

Methods

Expression vectors and antibodies

The plasmids pCDNA3–Rab5^{S34N}-Myc, pEGFP-4–Rab5a, pCMV–Rab5a-Myc and pCDNA3–Rab5^{Q79L}-Myc were gifts from M. Zerial. The plasmids pDCR-H–Ras^{V12}, pDCR-H–Ras^{V12,C40} and pDCR-H–Ras^{V12,S35} were gifts from J. Downward. PI(3)K^{CD2p110} was a gift from D. Cantrell. pCDNA3–Rac^{Q61L} and pCDNA–Rac^{N17} were gifts from S. Gutkind. pEGFP–RN-tre and pCDNA–HA1–RN-tre were described previously³. Haemagglutinin (HA)- or glutathione S-transferase (GST)-tagged RN-tre mutants were obtained by recombinant polymerase chain reaction (PCR) and subcloning in pCDNA–HA or pGEX-6P (Pharmacia). Human ACTN4 was cloned by PCR with reverse transcription (RT–PCR), into either pEGFP-C3 (Clontech) or pCDNA3.1–Flag (Invitrogen). All constructs were verified by sequence analysis. Primers used are available upon request.

An anti-ACTN4 rabbit serum was raised against the peptide MGDYMAQEDDW. Other antibodies were: anti-HA (12CA5; BABCO), anti-Myc (9E10; BABCO), anti-Rab5 (S-19; Santa Cruz Biotechnology), anti-Ras (pan-ras, Ab-1; Calbiochem), anti-Rac (Transduction Laboratories), anti-GST rabbit polyclonal, anti-actin mouse ascitic fluid (AC-40; Sigma) and an anti-RN-tre monoclonal antibody was generated in house.

Target sequences for RNAi were: for RN-tre, R1 5′-AACATCCTTATTCATGTGC-3′, R2 5′-AAAGCGGGCAAAGTAGACCT-3′; for ACTN4, A1 5′-AAAGCCCTCATTCGCAAGCAC-3′, A2 5′-AAGCAGTATGAACGCAGATC-3′. Control oligos were designed by introducing three mismatched nucleotides in R1 and in A1. Human fibroblasts (Detroit 551) were transfected twice at 24 h intervals with Oligofectamine (Invitrogen). The final concentration of the small interfering RNA duplex in the culture medium was 100 nM.

Immunofluorescence microscopy and biochemical assays

Immunofluorescence microscopy, immunoprecipitation and immunoblotting were performed as described^{12,28}. Cells were lysed on ice in a buffer containing 1% Triton X-100 (Pierce), 10 mM Tris-HCl at pH 7.6, 5 mM EDTA, 150 mM NaCl plus phosphatase and protease inhibitors (30 mM sodium pyrophosphate, 50 mM NaF, 100 μ M sodium orthovanadate, 2 mM phenyl methylsulphonylfluoride and 50 μ g ml^{−1} aprotinin). For the experiments shown in Fig. 3a, b, the buffer was supplemented with 0.1% SDS and 1% Na deoxycholate.

High-speed co-sedimentation assays were performed using 10 μ M of human non-muscle G-actin, polymerized in F-actin by addition of 50 mM KCl, 1 mM MgCl₂ and 1 mM ATP for 1 h at 30 °C. F-Actin was diluted to 2 μ M and then incubated with 0.4 μ M RN-tre–GST fragments for 10 min at room temperature. The samples were centrifuged at 150,000g in a Beckman TLA 120.2 at 4 °C for 1 h. Equal amounts of supernatant and pellet were analysed.

Unless otherwise indicated, ectopic expression of the various constructs was performed using the FuGENE 6 transfection reagent (Roche) in mouse embryo fibroblasts (MEF). Serum starvation (20 h) was initiated 10 h after transfection, followed by treatment with platelet derived growth factor (PDGF-BB) (10 ng ml^{−1}, 5 min, 37 °C); staining was performed immediately afterwards.

The quantitative experiment in Table 1 was performed by co-microinjection. Cells were serum starved for 16 h and then microinjected with the plasmids. PDGF treatment was performed 4 h after the microinjection, and cells were immediately stained with phalloidin. Coexpression of all constructs was verified in parallel samples, with double or triple stainings.

The macropinocytosis assay was performed as described²⁹.

Received 6 February; accepted 5 April 2004; doi:10.1038/nature02542.

1. Zerial, M. & McBride, H. Rab proteins as membrane organizers. *Nature Rev. Mol. Cell Biol.* **2**, 107–117 (2001).
2. Barbieri, M. A. *et al.* Epidermal growth factor and membrane trafficking. EGF receptor activation of endocytosis requires Rab5a. *J. Cell Biol.* **151**, 539–550 (2000).
3. Lanzetti, L. *et al.* The Eps8 protein coordinates EGF receptor signalling through Rac and trafficking through Rab5. *Nature* **408**, 374–377 (2000).
4. Swanson, J. & Watts, C. Macropinocytosis. *Trends Cell Biol.* **5**, 424–428 (1995).
5. Hall, A. Rho GTPases and the actin cytoskeleton. *Science* **279**, 509–514 (1998).
6. Bar-Sagi, D. & Hall, A. Ras and Rho GTPases: a family reunion. *Cell* **103**, 227–238 (2000).
7. Suetsugu, S., Yamazaki, D., Kurisu, S. & Takenawa, T. Differential roles of WAVE1 and WAVE2 in dorsal and peripheral ruffle formation for fibroblast cell migration. *Dev. Cell* **5**, 595–609 (2003).
8. Eriksson, A., Siegbahn, A., Westermark, B., Heldin, C. H. & Claesson-Welsh, L. PDGF α - and β -receptors activate unique and common signal transduction pathways. *EMBO J.* **11**, 543–550 (1992).
9. Bar-Sagi, D. & Feramisco, J. R. Induction of membrane ruffling and fluid-phase pinocytosis in quiescent fibroblasts by ras proteins. *Science* **233**, 1061–1068 (1986).
10. Ridley, A. J., Paterson, H. F., Johnston, C. L., Diekmann, D. & Hall, A. The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. *Cell* **70**, 401–410 (1992).
11. Dharmawardhane, S., Sanders, L. C., Martin, S. S., Daniels, R. H. & Bokoch, G. M. Localization of p21-activated kinase 1 (PAK1) to pinocytic vesicles and cortical actin structures in stimulated cells. *J. Cell Biol.* **138**, 1265–1278 (1997).
12. Scita, G. *et al.* EPS8 and E3B1 transduce signals from Ras to Rac. *Nature* **401**, 290–293 (1999).
13. Horiuchi, H. *et al.* A novel Rab5 GDP/GTP exchange factor complexed to Rabaptin-5 links nucleotide exchange to effector recruitment and function. *Cell* **90**, 1149–1159 (1997).
14. Tall, G. G., Barbieri, M. A., Stahl, P. D. & Horazdovsky, B. F. Ras-activated endocytosis is mediated by the Rab5 guanine nucleotide exchange activity of RIN1. *Dev. Cell* **1**, 73–82 (2001).

15. Spaargaren, M. & Bos, J. L. Rab5 induces Rac-independent lamellipodia formation and cell migration. *Mol. Biol. Cell* **10**, 3239–3250 (1999).
16. Benmerah, A., Bayrou, M., Cerf-Bensussan, N. & Dautry-Varsat, A. Inhibition of clathrin-coated pit assembly by an Eps15 mutant. *J. Cell Sci.* **112**, 1303–1311 (1999).
17. Joneson, T., White, M. A., Wigler, M. H. & Bar-Sagi, D. Stimulation of membrane ruffling and MAP kinase activation by distinct effectors of RAS. *Science* **271**, 810–812 (1996).
18. Wennstrom, S. *et al.* Activation of phosphoinositide 3-kinase is required for PDGF-stimulated membrane ruffling. *Curr. Biol.* **4**, 385–393 (1994).
19. Honda, K. *et al.* Actinin-4, a novel actin-bundling protein associated with cell motility and cancer invasion. *J. Cell Biol.* **140**, 1383–1393 (1998).
20. Djinnovic-Carugo, K., Young, P., Gautel, M. & Saraste, M. Structure of the alpha-actinin rod: molecular basis for cross-linking of actin filaments. *Cell* **98**, 537–546 (1999).
21. Araki, N., Hatae, T., Yamada, T. & Hirohashi, S. Actinin-4 is preferentially involved in circular ruffling and macropinocytosis in mouse macrophages: analysis by fluorescence ratio imaging. *J. Cell Sci.* **113**, 3329–3340 (2000).
22. Tocque, B. *et al.* Ras-GTPase activating protein (GAP): a putative effector for Ras. *Cell. Signal.* **9**, 153–158 (1997).
23. Greenwood, J. A., Theibert, A. B., Prestwich, G. D. & Murphy-Ullrich, J. E. Restructuring of focal adhesion plaques by PI 3-kinase. Regulation by PtdIns (3,4,5)-p(3) binding to alpha-actinin. *J. Cell Biol.* **150**, 627–642 (2000).
24. Corgan, A. M., Singleton, C., Santos, C. B. & Greenwood, J. A. Phosphoinositides differentially regulate alpha-actinin flexibility and function. *Biochem. J.* **378**, 1067–1072 (2004).
25. Fraley, T. S. *et al.* Phosphoinositide binding inhibits alpha-actinin bundling activity. *J. Biol. Chem.* **278**, 24039–24045 (2003).
26. Krueger, E. W., Orth, J. D., Cao, H. & McNiven, M. A. A dynamin-cortactin-Arp2/3 complex mediates actin reorganization in growth factor-stimulated cells. *Mol. Biol. Cell* **14**, 1085–1096 (2003).
27. Schafer, D. A. *et al.* Dynamin2 and cortactin regulate actin assembly and filament organization. *Curr. Biol.* **12**, 1852–1857 (2002).
28. Fazioli, F. *et al.* Eps8, a substrate for the epidermal growth factor receptor kinase, enhances EGF-dependent mitogenic signals. *EMBO J.* **12**, 3799–3808 (1993).
29. Seastone, D. J. *et al.* The WASp-like protein scar regulates macropinocytosis, phagocytosis and endosomal membrane flow in Dictyostelium. *J. Cell Sci.* **114**, 2673–2683 (2001).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements This work was supported by grants from the Associazione Italiana per la Ricerca sul Cancro and Human Science Frontier Program to P.P.D.F. and G.S., and by grants from the Association for International Cancer Research, European Community (VI Framework), Italian Ministry for University and Research and Consiglio Nazionale delle Ricerche to P.P.D.F. We thank E. Frittoli for technical assistance. L.L. was supported in part by a fellowship from the Associazione Sviluppo Piemonte.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.P.D.F. (difiore@ifom-firc.it).

A highly active synthetic mammalian retrotransposon

Jeffrey S. Han & Jef D. Boeke

Department of Molecular Biology and Genetics and High Throughput Biology Center, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

LINE-1 (L1) elements are retrotransposons that comprise large fractions of mammalian genomes¹. Transcription through L1 open reading frames is inefficient owing to an elongation defect², inhibiting the robust expression of L1 RNA and proteins, the substrate and enzyme(s) for retrotransposition^{3–5}. This elongation defect probably controls L1 transposition frequency in mammalian cells. Here we report bypassing this transcriptional defect by synthesizing the open reading frames of L1 from synthetic oligonucleotides, altering 24% of the nucleic acid sequence without changing the amino acid sequence. Such resynthesis led to greatly enhanced steady-state L1 RNA and protein levels. Remarkably, when the synthetic open reading frames were substituted for the wild-type open reading frames in an established retrotransposition assay⁴, transposition levels increased more than 200-fold. This indicates that there are

probably no large, rigidly conserved *cis*-acting nucleic acid sequences required for retrotransposition within L1 coding regions. These synthetic retrotransposons are also the most highly active L1 elements known so far and have potential as practical tools for manipulating mammalian genomes.

L1 retrotransposons account, directly or indirectly, for more than 30% of mammalian genomes by mass¹, by means of self-mobilization and trans-mobilization of *Alu* elements⁶. A full-length (about 6-kilobase) L1 (Fig. 1a) consists of two open reading frames, ORF1 and ORF2, that encode proteins required for retrotransposition^{3,4}. Although ORF1 translation is assumed to occur by 5' cap-binding and scanning, the mechanism for ORF2 translation initiation is unknown. ORF1 and ORF2 assemble into a ribonucleo-protein complex^{7,8} that enters the nucleus and nicks the target site, priming the reverse transcription⁵ of L1 RNA. L1 proteins show a strong preference for acting on the L1 RNA that encoded them, a phenomenon known as *cis*-preference^{9,10}. Nicking and reverse transcription activities are provided by ORF2 protein^{3,11}, but the function of ORF1 is unknown.

A highly active L1 element would be potentially useful as a tool for mammalian genetics. However, ORF1 and ORF2 sequences are poorly transcribed, and consequently minuscule amounts of functional, full-length L1 RNA are produced in mammalian expression systems². We reasoned that the retrotransposition frequency of L1 might be limited in part by this poor transcriptional elongation. To circumvent this problem, we designed a synthetic mouse ORF2 (smORF2) based on favoured codons in highly expressed mammalian genes. This altered 24% of the nucleic acid sequences but did

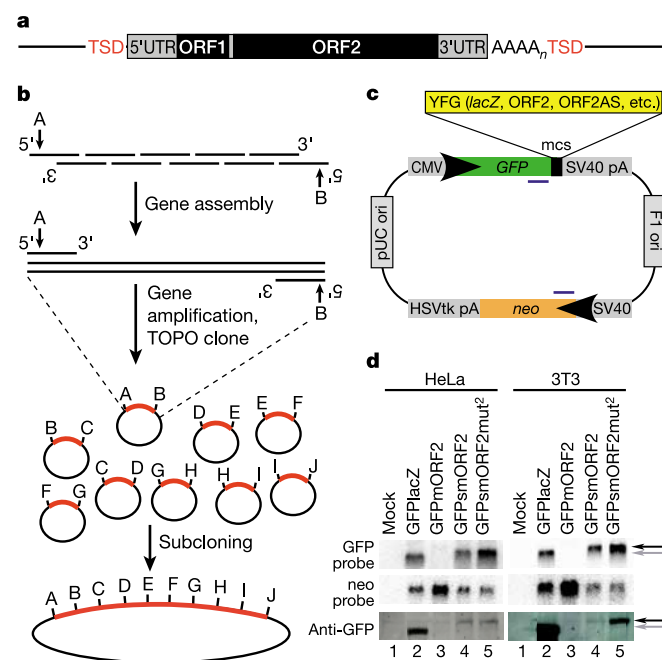


Figure 1 Synthesis and expression of synthetic mouse ORF2. **a**, L1 structure. TSD, target site duplication; UTR, untranslated region. **b**, Overview of gene synthesis. Oligonucleotides encoding each fragment were mixed in a PCR assembly reaction and subsequently used as template amplification. Amplification products were cloned and ligated together with unique restriction sites (labelled A to J). **c**, Plasmid structures. The test sequences (*lacZ*, mORF2 or smORF2) are fused, in frame, downstream of the GFP ORF. An independent *neo* transcript is used to monitor transfection efficiency and loading. Blue lines represent probes used in **d**. **d**, Analysis of smORF2 expression. Top, RNA expression of GFP, GFPmORF2 and GFPsmORF2. Middle, RNA expression of loading control. Bottom, protein expression of GFP, GFPmORF2 and GFPsmORF2. Grey and black arrows indicate the expected sizes of GFP and GFPmORF2/GFPsmORF2, respectively.

not change the protein (Supplementary Fig. S1). This was done in an attempt to destroy any *cis*-acting sequences responsible for poor transcription, including a previously described adenosine-rich bias that might be responsible for the transcription defect². The adenosine content of the smORF2 sequence was decreased to 26% (in comparison with 40% for native mouse ORF2 (mORF2)). Using oligonucleotide-based gene synthesis¹² (Fig. 1b), the recoded gene was synthesized in nine roughly 500-base-pair (bp) fragments and was subsequently assembled by ligation into smORF2. The expression of smORF2 was tested in a fusion vector with green fluorescent protein (GFP) (Fig. 1c) as described previously². In both

human and mouse cells, transfection of GFPsmORF2 led to a massive increase in RNA compared with wild-type GFPmORF2 (Fig. 1d, top panel, lanes 3 and 4). The introduction of two mutations that abolish the endonuclease and reverse transcriptase activities of mORF2 provided a further slight increase in smORF2 RNA levels (Fig. 1d, top panel, lanes 5), which is consistent with the known toxicity of ORF2 overexpression in other organisms^{5,11}. Probing for the vector-encoded *neo* transcript showed that these increases in RNA were not due to differences in transfection efficiency or loading (Fig. 1d, middle panel). Immunoblotting these samples with anti-GFP (Fig. 1d, bottom panel) showed that protein levels were correlated with RNA increase, marking the first instance of the reproducible expression of detectable amounts of recombinant full-length ORF2 protein in a mammalian system. However, GFPsmORF2 protein levels were still low relative to the control GFP Δ lacZ protein, suggesting that the ORF2 sequence might also be poorly translated or unstable.

We next sought to determine whether the increased RNA levels led to altered retrotransposition efficiency. We used an established tissue culture assay (Fig. 2a) to measure relative retrotransposition frequencies in HeLa cells. mORF2 was replaced with smORF2 in a full-length mouse L1 to make a partly synthetic mouse L1 (psmL1). Because we were concerned that recoded mORF2 might lack potentially important *cis*-acting sequences required for retrotransposition (for example, an internal ribosomal entry site), we also constructed a partly synthetic version of ORF2 (psmL1-2) in which the first roughly 500 bp of mORF2 consisted of wild-type L1 sequence and the remainder was synthetic. In HeLa cells, both psmL1 and psmL1-2 were about 20–25-fold more active than wild-type mL1 (Fig. 2b). Synthesis and incorporation of a synthetic mORF1 (smORF1) and partly synthetic mORF1 variants led to further increases in retrotransposition, reaching a maximum of more than 200-fold increase over wild type (Fig. 2b) in the element with two fully synthetic ORFs.

To verify that these smL1 G418-resistant colonies resulted from authentic L1 retrotransposition, we characterized six smL1 insertions. The mutant loci were identified by inverse polymerase chain reaction (PCR), enabling the amplification of each complete insertion and flanking sequence. For each primer pair, parental HeLa cells produced only empty site products (Fig. 3a, odd-numbered lanes), whereas the respective G418-resistant clones produced both empty site and filled smL1 insertion products of predicted sizes (Fig. 3a, even-numbered lanes). Amplicons were cloned and sequenced to determine their general structures and genomic flanks, summarized in Fig. 3b. All showed a properly spliced *neo* gene, a poly(A) tail, and most (five of six) had target site duplications 5–108 bp long. Insertion no. 10 had a 10-bp target deletion and insertion no. 18 had a 5' L1 inversion, features commonly found in L1 insertions^{13–15}. In addition, various chromosomes served as targets, and the endonuclease cleavage sites inferred from target site duplications matched the previously reported degenerate consensus (5'-TTTT/AA-3' on the bottom strand)^{3,16} (Fig. 3c). Taken together, these results suggest that the synthetic L1 retrotransposes

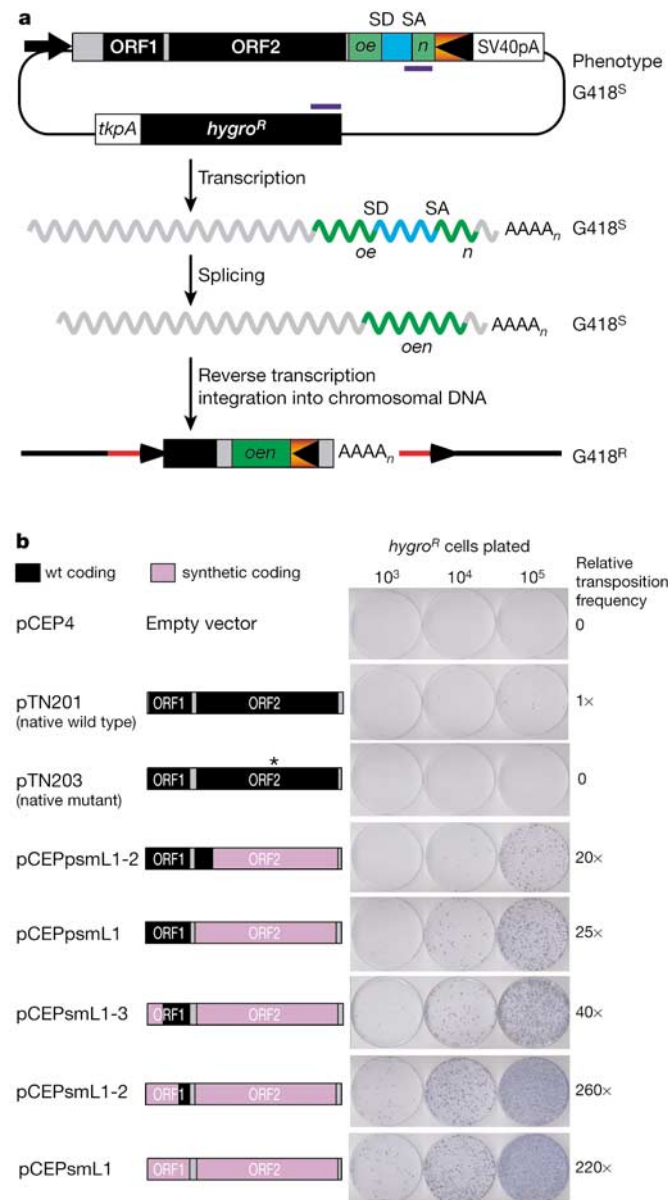


Figure 2 Retrotransposition of synthetic mL1. **a**, The retrotransposition assay. The L1 element contains an intron-interrupted *neo* reporter in the 3' untranslated region with its own promoter and polyadenylation signal. Only when *neo* is transcribed from the L1 promoter, spliced, reverse transcribed and integrated into the genome does a cell become G418-resistant⁴. Blue lines represent probes for RNA analysis (Fig. 4). SD, splice donor; SA, splice acceptor. **b**, Retrotransposition was assayed in HeLa cells ($N = 3$). pTN201 contains only wild-type native mouse L1 sequence, and pTN203 contains wild-type native mouse L1 sequence with a D709Y reverse transcriptase point mutation²². The average absolute number of colonies for pTN201 was 440 events per 10^6 transfected cells.

Table 1 High-frequency retrotransposition in mouse cells

Plasmid	Relative transposition frequency		
	HeLa	3T3	L
pCEP4 (empty vector)	0	0	0
pTN201 (native mouse wild type)	<0.005	<0.002	<0.002
pTN203 (native mouse mutant)	0	0	0
pJM101L1 (native human wild type)	0.13	0.017	0.07
pCEPsmL1 (synthetic mouse wild type)	1	1	1
pCEPsmL1mut ² (synthetic mouse mutant)	0	0	<0.002

With the use of the transient assay¹⁷, synthetic mouse L1 (pCEPsmL1) retrotransposition frequency was compared with that of wild-type native human L1 and wild-type native mouse L1 ($N = 3$). The average absolute numbers of colonies per pJM101L1rp (colonies per 10^6 transfected cells) for HeLa, 3T3 and L cells were 2,904, 108 and 1,568, respectively.

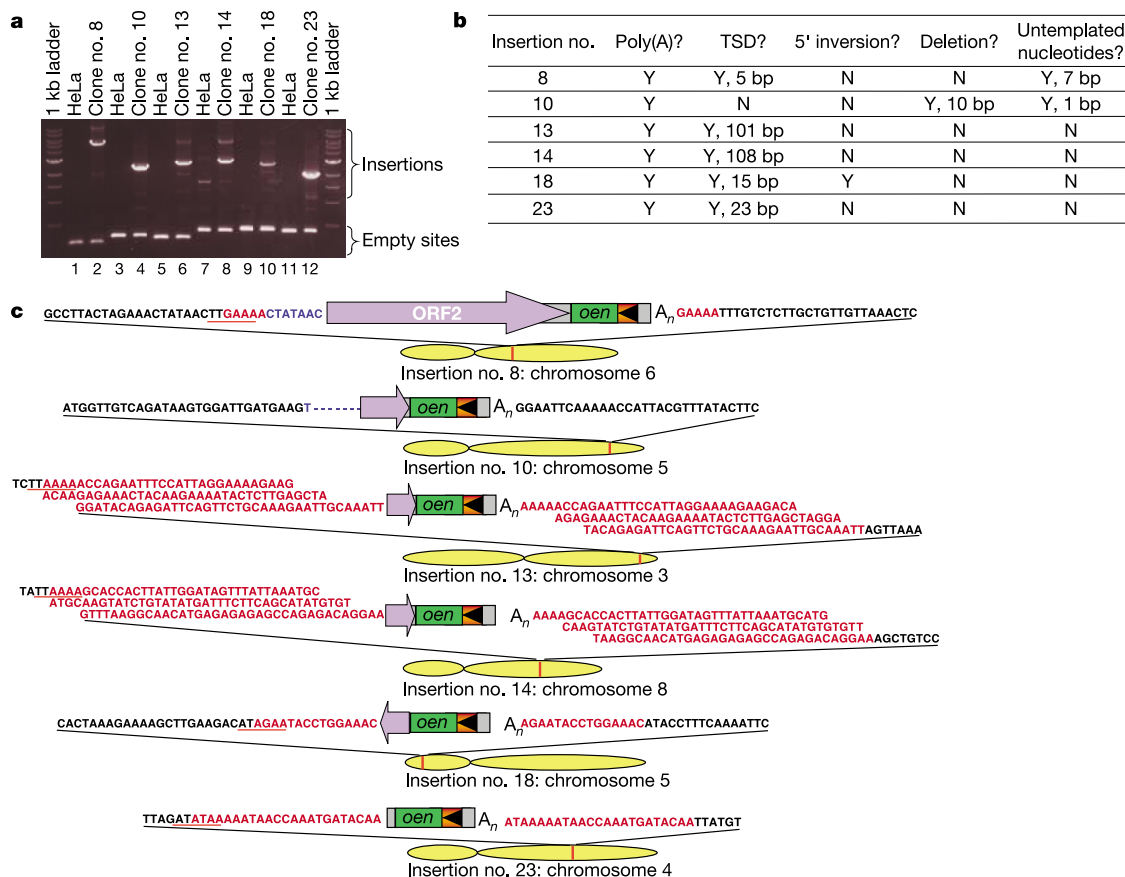


Figure 3 Synthetic mouse L1 uses the standard retrotransposition mechanism. **a**, Primers flanking each insertion were used for amplification from G418-resistant clones (see the text). **b**, Characteristics of cloned insertions. TSD, target site duplication. **c**, Structure and flanking sequence of cloned insertions are shown. Insertion no. 8 contained an additional 7 bp (highlighted in blue) not found in the human genome

sequence (<http://genome.ucsc.edu>). Insertion no. 10 contained one untemplated base pair relative to the human genome sequence database followed by a 10-bp deletion (indicated in blue) immediately upstream of the L1 insertion. TSDs are highlighted in red, and presumptive endonuclease cleavage sites are underlined.

by target-primed reverse transcription and not some aberrant mechanism.

We also compared the activity of the synthetic mouse L1 retrotransposons with wild-type human and mouse L1 in mouse cells. Because episomal plasmids used to introduce marked retrotransposons do not replicate efficiently in mouse cells, we used a transient retrotransposition assay¹⁷ in 3T3 and L cells. We also performed the transient assay in HeLa cells, verifying the relative retrotransposition frequencies obtained with the standard assay (compare pTN201 and pCEPsmL1 from Fig. 2b and Table 1).

The synthetic mouse L1 (pCEPsmL1) underwent retrotransposition at much higher frequencies (more than 200-fold) than its wild-type counterpart in mouse cells. In addition, we compared smL1 with a human L1 (pJM101L1rp) because L1rp has previously been used to generate transgenic mouse lines and thus serves as a benchmark for retrotransposition frequencies in mice. smL1 was significantly more active than L1rp in all cell types tested, making it the most active L1 element known so far. Introducing catalytic mutations into smL1 to produce smL1mut² essentially abolished retrotransposition, confirming that smL1 retrotransposition depends on ORF2 endonuclease and reverse transcriptase functions.

Northern blot analysis of wild-type full-length mL1 and its synthetic counterparts revealed that increasing lengths of synthetic L1 sequence led to increasing full-length L1 RNA levels (Fig. 4). pCEPsmL1mut² was used in place of pCEPsmL1 because pCEPsmL1 was difficult to maintain episomally, as determined by the *hygro* transfection/loading control (data not shown). This

suggests, because the intact pCEPsmL1 plasmid is not maintained in transfected cells for long periods, that the reported increased retrotransposition frequencies of smL1 relative to native mouse or human L1 are underestimates. In addition, the increased RNA levels suggest that increased RNA expression is, at least in part, responsible for enhanced retrotransposition levels. However, because codon usage was optimized for mammalian cells, improved translational efficiency might also be significant. We also cannot rule out the possibility that recoding destroyed regulatory nucleic acid motifs that affect retrotransposition in multiple ways.

It is already known that both the 5' and 3' untranslated regions and the interORF region of L1 are not required for retrotransposition (refs 4, 18, and R. S. Alisch and J. V. Moran, personal communication). This indicates that if any essential nucleic acid

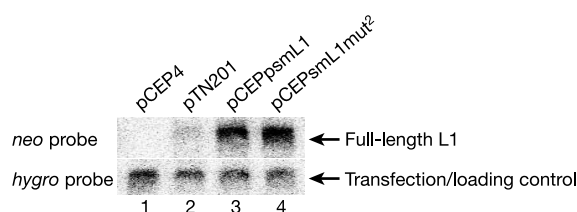


Figure 4 High-frequency retrotransposition in mouse cells: total RNA analysis of smL1 expression. Expression of native, partly synthetic, and completely synthetic mL1 was compared in HeLa cells.

motifs exist in L1, they are within the coding regions. Because our synthetic mouse L1 is extensively mutagenized throughout the coding regions but still retrotransposes with startlingly high efficiency, either L1 essential nucleic acid sequences are small and fortuitously preserved in smL1, or L1 essential nucleic acid sequences are highly tolerant to mutations. It is also plausible that, aside from encoding functional proteins, there are no special requirements for L1 nucleic acid sequence. Only further investigation will distinguish between these possibilities, but it is of interest to note that an absence of required nucleic acid motifs might necessitate unconventional explanations for mysterious aspects of the L1 life cycle such as the ORF2 translational initiation mechanism and *cis*-preference.

Finally, our synthetic L1 might represent a major step forward in efforts to design a useful random mutagenesis system in mice. Transposons are useful genetic tools because of their ability to produce mutations by adding new DNA sequence, 'tagging' the disrupted gene for easy cloning and identification¹⁹. However, in mammalian systems the low frequency of retrotransposition has precluded the practical large-scale use of retrotransposons for mutagenesis. For example, the theoretical potential of L1 as a mutagenic agent has been shown in mice, but so far the frequency of progeny carrying a new insertion has reached a maximum of less than 10%²⁰. An ideal mutagenesis system would produce multiple new insertions per progeny animal such that each carries a new mutant gene. Thus, the retrotransposition rate of L1 in current mouse models is one to two orders of magnitude lower than desired. In mouse cells our synthetic L1 retrotransposes at a frequency ranging from 15-fold to 50-fold higher than L1rp. If this increase in L1 activity in mouse cells translates to animals, we envisage creating transgenic mouse lines that, once made, continuously generate easily clonable random mutations in each progeny animal, without embryonic stem cell manipulations. Such a line would be extremely useful for practical forward genetic screens. In addition, cataloguing and storing mutants (or mutant sperm) could allow the comprehensive knockout of mouse genes. □

Methods

Gene synthesis

smORF2 and ORF1 sequences were created by replacing each codon in the mouse L1 ORFs with the favoured codons in highly expressed human genes²¹ (Supplementary Table S2). The sequence was further altered with silent mutations introducing unique cleavage sites and eliminating potential hairpins that might have inhibited gene assembly. 60-mer oligonucleotides collectively encoding both strands of smORF2 were ordered from Qiagen, and gene synthesis¹² was performed on each ~500-bp segment as shown in Fig. 1b. Assembly reactions contained each primer at 30 nM and 1 × ExTaq mix (Takara) in a total of 25 µl. Amplification reactions contained each outer primer at 0.5 µM, 2.5 µl assembly reaction, and 1 × ExTaq mix in a total volume of 25 µl. PCR conditions were 94 °C for 4 min, 25 cycles of 94 °C for 30 s, 65 °C for 30 s, and 72 °C for 30 s, followed by 72 °C for 7 min. PCR products were cloned into pCRII with the TOPO-TA cloning kit (Invitrogen). A total of 24–48 clones were sequenced for each fragment and mutations were removed by standard cloning techniques. Finally, synthesized fragments were ligated together in pBluescriptKS⁺. Oligonucleotide sequences used are shown in Supplementary Table S1.

Plasmids

pGFP_{lacZ} and pGFP_{mORF2} are described elsewhere². pTN201 and pTN203 were gifts from H. Kazazian²². pJM101L1rp was provided by J. Moran¹⁷. Detailed descriptions of the construction of the plasmids are available from the authors on request.

Cell culture and transfection

HeLa cells, 3T3 cells and L cells were gifts from the laboratories of J. Moran, S. Desiderio and J. Nathans. Cells were grown in DMEM medium (Invitrogen) with 10% FBS (Invitrogen) and 1% penicillin/streptomycin (Invitrogen).

Transfections were performed with Eugene 6 (Roche) in six-well dishes. The transfection mix consisted of 100 µl Opti-MEM (Invitrogen), 3 µl Eugene and 2 µg DNA. For downstream northern or immunoblot analyses, cells were harvested 36–48 h after transfection.

Northern blot analysis

Total RNA was isolated with TRIzol reagent (Invitrogen) in accordance with the manufacturer's instructions. Total RNA (6 µg) from each sample was treated with 10 units of DNase I for 15 min at 37 °C, then run on a 0.8% agarose/formaldehyde gel, blotted overnight to a Genescreen plus nylon membrane (NEN) in 10 × SSC, and crosslinked by

ultraviolet radiation. Prehybridizations and hybridizations were both performed in ULTRAhyb (Ambion) at 42 °C. The following [γ -³²P]ATP end-labelled oligonucleotides were used as probes: GFP probe, JB4057; GFP plasmid *neo* probe, JB4059; transposition plasmid *neo* probe, JB4541; *hyg* probe, JB6341. Washes were performed in 2 × SSC, 0.1% SDS and in 0.2 × SSC, 0.1% SDS. Radioactive signal was detected with Fuji imaging plates and a Fuji scanner (BAS-1500). For subsequent reprobing, membranes were stripped with three 10-min washes in boiling 0.1 × SSC, 1% SDS.

Immunoblot analysis

Cells were harvested in 5% SDS/PBS; this was followed by sonication. Total lysates were subjected to 7.5% SDS-polyacrylamide-gel electrophoresis and transferred to poly(vinylidene difluoride) (Amersham). Antibody incubations were performed in PBS containing 0.05% Tween-20 and 5% milk. Washes were performed in PBS, 0.1% Tween-20. Anti-GFP(FL) antibody (Santa Cruz) was used at 1:250 dilution. Anti-rabbit IgG (Amersham) was used at 1:5,000 dilution. Blots were developed with ECL-plus (Amersham).

Retrotransposition assays

The standard retrotransposition assay in HeLa cells was performed essentially as described⁴. Transfected cells were selected with 200 µg ml⁻¹ hygromycin for 10–12 days, then counted and seeded in 600 µg ml⁻¹ G418 for 10 days. Colonies were stained with 0.4% Giemsa in PBS.

The transient retrotransposition assays in HeLa, 3T3 and L cells were performed essentially as described¹⁷. Each transposition construct was cotransfected with the GFP-expressing plasmid pTracerEF (Invitrogen) to normalize for transfection efficiency. At 24 h after transfection, cells were split 1:2, 1:20 and 1:200 into 100-mm dishes. At 36 h after transfection, the diluted cells were selected with G418 and the remaining cells were analysed for GFP expression by flow cytometry to normalize for transfection efficiency. 3T3 cells were selected in 1 mg ml⁻¹ G418; L cells were selected in 400 µg ml⁻¹ G418. Colonies were stained with 0.4% Giemsa or 0.5% Coomassie brilliant blue.

Cloning of retrotransposition events

Integration sites were determined by inverted PCR essentially as described²³. Genomic DNA (5 µg) from each clone was digested with *Eco*RI, inactivated by heat, diluted to 1 ml and ligated overnight, precipitated with ethanol, resuspended in 30 µl water and subjected to two rounds of inverted PCR with oligos JB6466/JB6467 (round 1) and JB6468/JB6469 (round 2). Sequencing with JB3529, JB3530 and JB3531 identified the 3' flanking sequences. Primers based on flanking sequence were used to amplify intact smL1 insertions, which were subsequently sequenced.

Received 22 January; accepted 30 March 2004; doi:10.1038/nature02535.

- Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
- Han, J. S., Szak, S. T. & Boeke, J. D. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* **429**, 268–274 (2004).
- Feng, Q., Moran, J. V., Kazazian, H. H. Jr & Boeke, J. D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916 (1996).
- Moran, J. V. *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917–927 (1996).
- Cost, G. J., Feng, Q., Jacquier, A. & Boeke, J. D. Human L1 element target-primed reverse transcription *in vitro*. *EMBO J.* **21**, 5899–5910 (2002).
- Dewannieux, M., Esnault, C. & Heidmann, T. LINE-mediated retrotransposition of marked Alu sequences. *Nature Genet.* **35**, 41–48 (2003).
- Martin, S. L. Characterization of a LINE-1 cDNA that originated from RNA present in the ribonucleoprotein particles: implications for the structure of an active mouse LINE-1. *Gene* **153**, 261–266 (1995).
- Koloshva, V. O. & Martin, S. L. *In vitro* properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. *Proc. Natl Acad. Sci. USA* **94**, 10155–10160 (1997).
- Esnault, C., Maestre, J. & Heidmann, T. Human LINE retrotransposons generate processed pseudogenes. *Nature Genet.* **24**, 363–367 (2000).
- Wei, W. *et al.* Human L1 retrotransposition: *cis* preference versus *trans* complementation. *Mol. Cell Biol.* **21**, 1429–1439 (2001).
- Mathias, S. L., Scott, A. F., Kazazian, H. H. Jr, Boeke, J. D. & Gabriel, A. Reverse transcriptase encoded by a human transposable element. *Science* **254**, 1808–1810 (1991).
- Stemmer, W. P., Crameri, A., Ha, K. D., Brennan, T. M. & Heyneker, H. L. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxynucleotides. *Gene* **164**, 49–53 (1995).
- Boissinot, S., Chevret, P. & Furano, A. V. L1 (LINE-1) retrotransposon evolution and amplification in recent human history. *Mol. Biol. Evol.* **17**, 915–928 (2000).
- Gilbert, N., Lutz-Prigge, S. & Moran, J. V. Genomic deletions created upon LINE-1 retrotransposition. *Cell* **110**, 315–325 (2002).
- Symer, D. E. *et al.* Human L1 retrotransposition is associated with genetic instability *in vivo*. *Cell* **110**, 327–338 (2002).
- Cost, G. J. & Boeke, J. D. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochemistry* **37**, 18081–18093 (1998).
- Wei, W., Morrish, T. A., Alisch, R. S. & Moran, J. V. A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events. *Anal. Biochem.* **284**, 435–438 (2000).
- Moran, J. V., DeBerardinis, R. J. & Kazazian, H. H. Jr Exon shuffling by L1 retrotransposition. *Science* **283**, 1530–1534 (1999).
- Hamer, L., DeZwaan, T. M., Montenegro-Chamorro, M. V., Frank, S. A. & Hamer, J. E. Recent advances in large-scale transposon mutagenesis. *Curr. Opin. Chem. Biol.* **5**, 67–72 (2001).
- Ostergaard, E. M. *et al.* A mouse model of human L1 retrotransposition. *Nature Genet.* **32**, 655–660 (2002).
- Haas, J., Park, E.-C. & Seed, B. Codon usage limitation in the expression of HIV-1 envelope glycoprotein. *Curr. Biol.* **6**, 315–324 (1996).

22. Naas, T. P. *et al.* An actively retrotransposing, novel subfamily of mouse L1 elements. *EMBO J.* **17**, 590–597 (1998).
23. Morrish, T. A. *et al.* DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nature Genet.* **31**, 159–165 (2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank Boeke laboratory members, especially Y. Aizawa, for helpful discussions and critical reading of the manuscript. We thank A. Yetil for assistance. This work was supported by the NIH (J.D.B.) and the Medical Scientist Training Program (J.S.H.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.D.B. (jboeke@jhmi.edu).

Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain

Jin-Biao Ma^{*}, Keqiong Ye^{*} & Dinshaw J. Patel

Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York 10021, USA

^{*} These authors contributed equally to this work

Short RNAs mediate gene silencing, a process associated with virus resistance, developmental control and heterochromatin formation in eukaryotes^{1–5}. RNA silencing is initiated through Dicer-mediated processing of double-stranded RNA into small interfering RNA (siRNA)^{6,7}. The siRNA guide strand associates with the Argonaute protein in silencing effector complexes, recognizes complementary sequences and targets them for silencing^{8–11}. The PAZ domain is an RNA-binding module found in Argonaute and some Dicer proteins and its structure has been determined in the free state^{12–14}. Here, we report the 2.6 Å crystal structure of the PAZ domain from human Argonaute eIF2c1 bound to both ends of a 9-mer siRNA-like duplex. In a sequence-independent manner, PAZ anchors the 2-nucleotide 3′ overhang of the siRNA-like duplex within a highly conserved binding pocket, and secures the duplex by binding the 7-nucleotide phosphodiester backbone of the overhang-containing strand and capping the 5′-terminal residue of the complementary strand. On the basis of the structure and on binding assays, we propose that PAZ might serve as an siRNA-end-binding module for siRNA transfer in the RNA silencing pathway, and as an anchoring site for the 3′ end of guide RNA within silencing effector complexes.

siRNA is a 19–23-base-paired (bp) duplex with 2-nucleotide (nt) 3′ overhangs at both ends, containing 5′-phosphates and 3′-hydroxyls^{7,15}. siRNA is not merely a consequence of Dicer processing, as both the length and ends are important for mediating RNA interference (RNAi)^{7,10,15}, whereby target messenger RNA is sequence-specifically degraded in an siRNA-programmed effector complex termed the RNA-induced silencing complex (RISC)^{8,9,16}. This suggests that specific structural features of siRNA constitute recognition targets for the protein components within the RNAi machinery. The PAZ domain can bind to single-stranded (ss) RNAs and siRNA duplexes^{12–14}, and requires the 2-nt 3′ overhang for efficient complex formation^{13,14}. We have determined a co-crystal structure of PAZ domain and a 9-mer RNA (5′-CGUGACUCU-3′) in order to illustrate the details of PAZ–RNA interaction and gain functional insights into the recognition process.

The structure of the PAZ–RNA complex reveals an unanticipated arrangement in which the 9-mer RNA, initially designed as a ssRNA ligand, forms a self-complementary siRNA-like A-form duplex, which is bound by the PAZ domain at each end (Fig. 1). The 3′-shifted pairing of each strand results in 2-nt, single-stranded 3′ overhangs at the duplex ends, which are characteristic of siRNA architecture. However, the short 7-bp duplex is unstable, due to the presence of three non-canonical pairs (Fig. 1a), raising the concern that the observed duplex formation could result from crystal-packing interactions. This is not the case, as the PAZ domain and the 9-mer RNA form a stable 2:2 complex in solution, as judged by gel filtration (Supplementary Fig. S1). Thus, PAZ binding apparently stabilizes siRNA-like duplex formation. Moreover, the absence of contacts between the two PAZ domains in the complex indicates independent binding of PAZ domains to the siRNA-like duplex, and suggests that such an arrangement should also hold for complex formation to the ends of a typical 21-nt siRNA.

The PAZ domain in the complex adopts a heart-shaped globular topology (Fig. 2a), with a twisted β-barrel consisting of six β-strands (β1–β3, β6–β8), capped by two amino-terminal α-helices (α1, α2) on one side and connected to an αβ module (β4–β5–α3) on the other side. The two strands of the siRNA-like duplex interact with a specific PAZ domain in a highly asymmetric manner. The strand bound with its 3′ end contacts the PAZ domain along its full 9-nt length, whereas the complementary strand makes contacts only with the 5′-terminal residue. The first 7 nucleotides constituting the RNA duplex structure, proceeding in the 5′ to 3′ direction, interact with the protein's carboxy-terminal tail and the positively charged surface formed by the strands β2, β3 and the β6–β7 loop (Fig. 2b). The bound RNA adopts a stacked 5′ to 3′ helical trajectory, except for a sharp turn (clockwise rotation of ~110° along the helical axis) in the phosphodiester backbone between the duplex and 2-nt 3′-overhang segments, thereby inserting the 2-nt ends into a central protein pocket formed between the barrel and αβ module. The same central RNA-binding pocket has been proposed previously from an evaluation of NMR chemical shift perturbations, mutagenesis and sequence conservation analysis^{12–14}. The free^{12–14} and RNA-bound

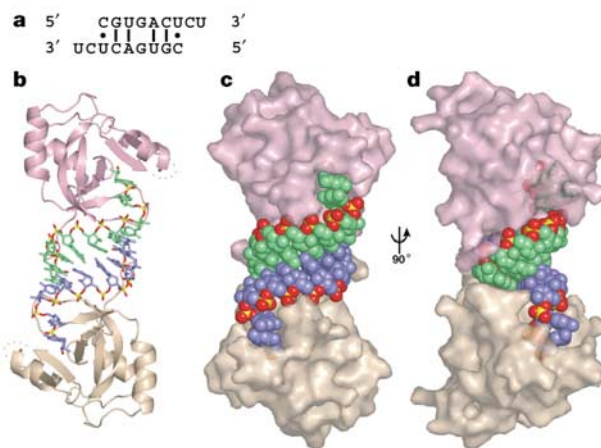


Figure 1 Overview of the PAZ–siRNA-like duplex structure **a**, The self-complementary siRNA-like duplex formed in the crystal. **b**, The entire complex of two PAZ domains bound to each end of an siRNA-like duplex. Protein and RNA are presented in ribbon and stick representations, respectively. **c**, The same view as **b**, but protein and RNA are presented in semi-transparent surface- and space-filling representations, respectively. **d**, 90° rotation of **c**. Note that PAZ domains predominantly contact the 3′-overhang-containing strands and dock with the 5′ ends of the complementary strands. The 2-nt 3′ overhangs are anchored within the pockets, with their base edges facing outwards towards solvent. The PAZ domains are coloured pink and beige and the RNA strands are blue and green, except for the phosphate groups of the RNA strands, in which phosphate is yellow and oxygen is red.

PAZ structures superimpose well around the RNA-binding pocket (Supplementary Fig. S2), indicating that the pocket is largely preformed.

The elongated pocket, which is lined by a large number of aromatic and hydrophobic residues, is generated by helix $\alpha 3$ and strand $\beta 4$ along one face, and strand $\beta 7$ and loop $\beta 2$ – $\beta 3$ along the opposite face (Fig. 2a). The two nucleotides at the 3' end retain a stacked A-form conformation (Fig. 3a), with the 3'-terminal residue buried deep within the pocket and targeted by numerous protein interactions (Fig. 3b). The non-bridging oxygens of the phosphodiester backbone linking the last two residues make four hydrogen bonds with Y309, Y314 and H269 and a water molecule stabilized by Y277. The sugar ring of the terminal residue is further anchored in place by van der Waals packing against L337 and T335, whereas its 2'- and 3'-hydroxyl groups form hydrogen bonds to the backbone amide and carbonyl groups of Y336, respectively. In addition, portions of the sugar ring and base of the terminal nucleotide are stacked over aromatic residue F292.

Among residues recognizing the 3'-terminal nucleotide, F292, Y309 and L337 are invariant, and H269, Y277, K313 and Y314 are conserved among PAZ homologues (Supplementary Fig. S3). These residues seem to be critical for efficient binding, as single mutations F292A and L337A cause a 615-fold and 155-fold reduction in binding affinity, and double mutation Y309F/Y314F causes a 46-fold reduction (Fig. 4a). Other conserved residues surrounding the pocket, such as V306, F310, P338 and V341, although not directly contacting the RNA, probably maintain the hydrophobic environment of the pocket. The strong binding of the 3'-terminal nucleotide, as reflected in numerous intermolecular contacts, explains the previous observation that single nucleotide uridine 5' monophosphate (UMP) can also bind in the same pocket¹². By contrast, the penultimate nucleotide from the 3' end makes only minimal contacts with PAZ—that is, a hydrogen bond between its 2'-hydroxyl group and Y277, and an electrostatic interaction with K313—and is largely exposed to solvent.

The base edges of the 2-nt 3'-overhang residues are facing outwards towards solvent and have no contact with the protein. The absence of specific hydrogen-bonding recognition of base edges as well as the available space along the base edges suggest that the pocket could accommodate 3'-overhang segments containing all nucleotide combinations. This feature is consistent with the potential function of PAZ, which is to recognize siRNAs of random sequence.

The PAZ domain laterally secures the siRNA-like duplex through hydrogen-bond and electrostatic interactions between five phosphodiester groups of the strand bound at the 3' end, and the basic residues R275, K333, R278, K264 and the C-terminal residues (Fig. 3c, d). The side chain of R278, which is buttressed in place by G262, forms two hydrogen bonds with adjacent phosphodiester groups (Fig. 3d). The importance of this recognition is manifested in the 17-fold loss in binding affinity observed for the R278A mutant (Fig. 4a).

No specific interactions are directed to the 2'-OH groups, which line the minor groove of the RNA helix, consistent with RNAi not generally being affected by 2'-OH group modifications within the siRNA duplex segment¹⁷. Nevertheless, the observed contacts between the PAZ domain and adjacent non-bridging phosphate oxygens (Fig. 3d) may provide a mechanism for discriminating the interphosphate separations, which are different between RNA and DNA helices. The PAZ domain also does not contact base edges within either groove of the RNA helix, indicative of non-sequence-specific molecular recognition.

The 5' end of the complementary strand is capped by loop $\beta 2$ – $\beta 3$ centred about residue M273 (Fig. 3c). This docking constitutes the only interaction involving the complementary strand, burying 100 Å² of the protein surface, in contrast to 900 Å² of total buried protein surface, upon complex formation. The aliphatic side chain of M273 extends towards the U7 base (Fig. 3c), but seems not to contribute to RNA binding, as no loss of binding was observed for mutant M273A (Fig. 4a). Furthermore, the PAZ domain does not contact the 5'-OH (and presumably 5'-phosphate, were it to be present) of the complementary strand.

To define more accurately the RNA structural determinants for PAZ recognition, we measured apparent dissociation constants (K_d) for RNA ligands with variant 3' ends, using surface plasmon resonance (Fig. 4b). RNAs containing both UU and AG 3' overhangs bind PAZ with the highest affinities (K_d values of 0.89–2.18 nM) among all tested constructs (Fig. 4b). The overhang contributes critically to the binding, as removal of one or two 3'-nucleotides reduces the binding affinity by 85- or >5,000-fold, consistent with the crystal structure of the complex, in which the PAZ domain optimally fits the duplex end with 2-nt 3' overhang in a sequence-independent manner.

In addition, the pocket favours the 2-nt overhang as ribonucleotides, because replacement with deoxyribonucleotides, which eliminates the two hydrogen bonds to the 2'-OH groups of the overhang

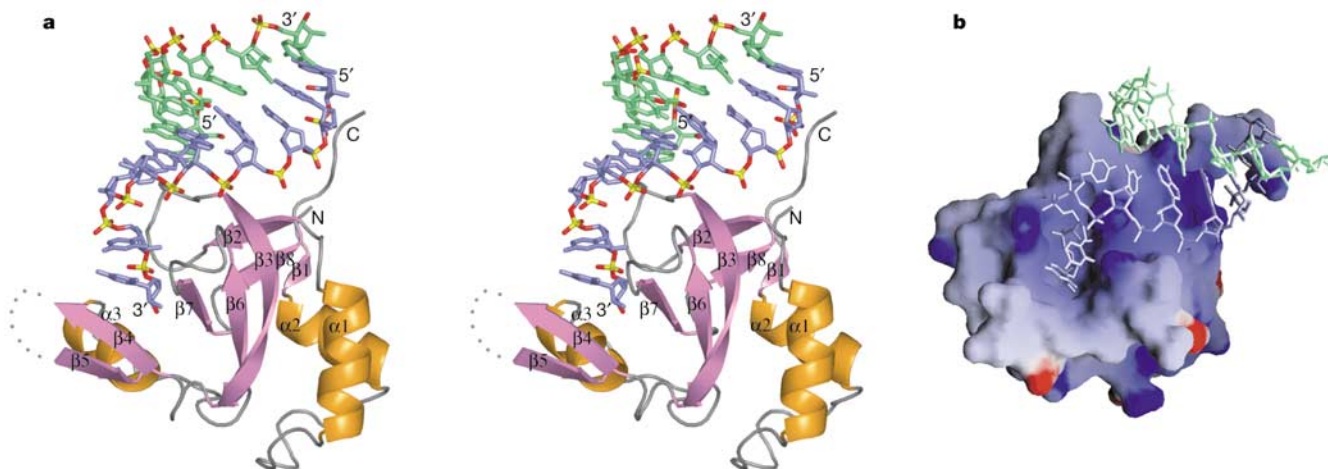


Figure 2 The PAZ-siRNA-like duplex subcomplex. **a**, Stereo view showing one PAZ domain interacting with an siRNA-like end. Relative to this PAZ domain, the RNA strand bound by the 3' end is coloured blue, whereas the strand bound by the 5' end is coloured

green. The protein secondary structure elements are labelled. **b**, Electrostatic surface of PAZ, with blue and red colours corresponding to positively and negatively charged patches, respectively.

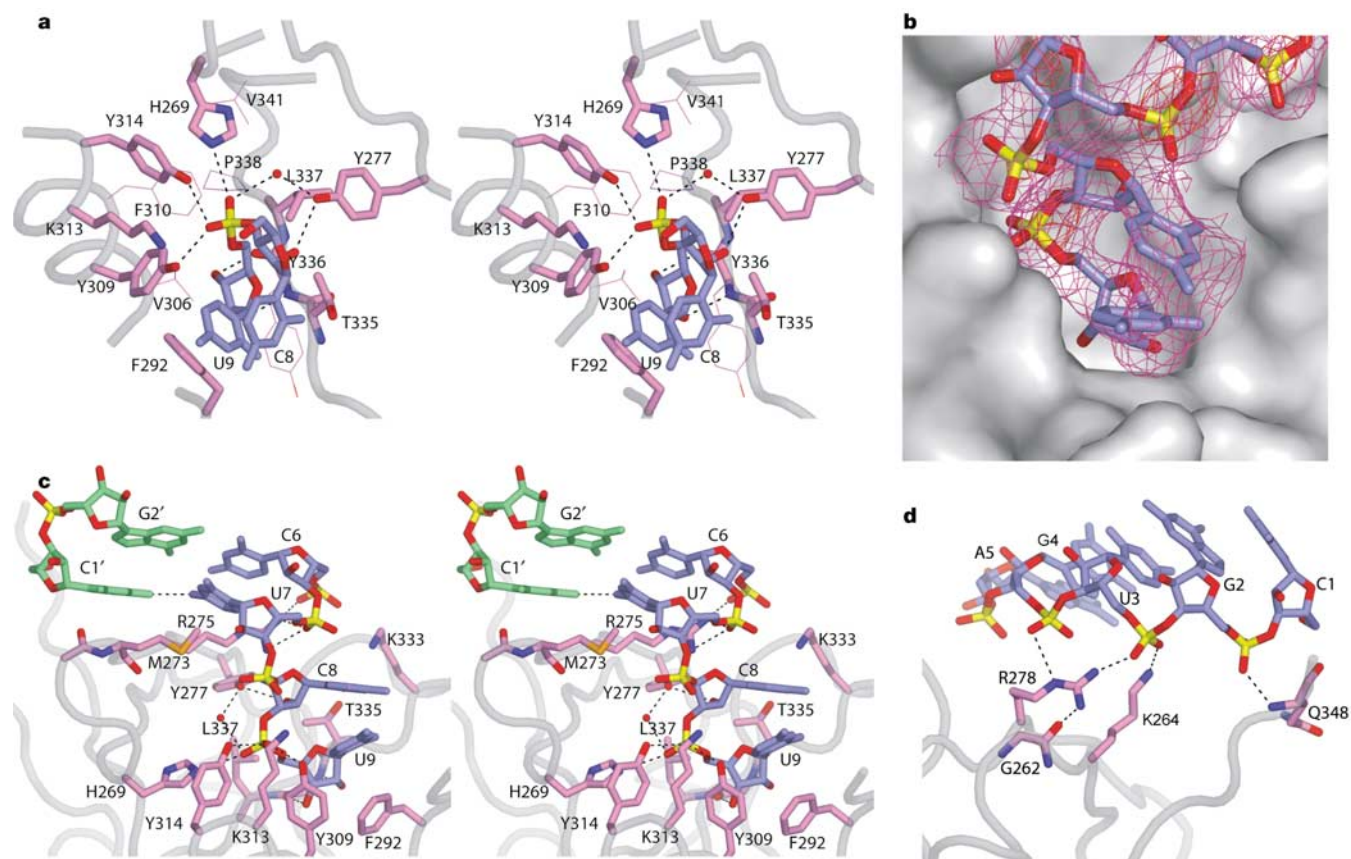


Figure 3 The details of PAZ-siRNA-like duplex interaction. **a**, The 2-nt 3'-overhang segment positioned within the conserved pocket. The residues directly contacting RNA are shown in a stick representation, whereas other residues contributing to the hydrophobic pocket are shown as thin lines. **b**, The RNA in stick representation positioned in the PAZ pocket in a surface representation. Also shown for RNA is the experimental density-modified electron density, contoured at 1.0 σ (magenta) and 4.5 σ (red). The contour of highest level is strongly indicative of the phosphate groups. **c**, Interaction of PAZ with the

duplex terminus and 3'-overhang segments, viewed from the left side of **a**. Note the change in the trajectory of RNA phosphodiester backbone between duplex and overhang segments. **d**, Interactions involving the RNA duplex segment. Potential hydrogen-bonding interactions are shown by dashed lines. **a** and **c** are shown in stereo view. Phosphates are coloured in yellow, oxygens in red, nitrogens in blue, protein side chains in violet, the 3'-end-bound RNA strand in blue and the 5'-end-bound RNA strand in green.

segment, reduces the binding by 100-fold (Fig. 4c). The pocket is incompatible with bulky 3'-end modifications, such as propanediol, fluorescein and puromycin, which decrease affinity by $>10^4$ -fold (Fig. 4c), but is fairly tolerant towards accommodation of a smaller 2'-*o*-methyl group at the terminal nucleotide (Fig. 4c).

The PAZ domain does not offer a strong binding site for the RNA 5' end, because elongation of the 3' overhang to 4 nt, which would disrupt the protein interaction with the 5'-terminal residue of the complementary strand, only slightly decreases the binding affinity by 5-fold (Fig. 4b). No binding occurs for ssRNA with a free 5' end (Fig. 4c), reinforcing this conclusion.

The PAZ domain is able to bind with the 10-nt overhang, in essence an ssRNA with a free 3' end, with a K_d of ~ 100 nM (Fig. 4b). Because of the predominant interaction of PAZ with the 3'-end-bound strand, the ssRNA might conceivably bind the pocket of PAZ by its 3' end and adopt an extended conformation along the β -barrel surface. The 50-fold reduction in affinity for ssRNA (10-nt overhang), relative to the siRNA end (2-nt overhang), may result from the lack of interaction with the 5'-terminal residue of the complementary strand and the increased flexibility of ssRNA due to the absence of duplex pairing. Together, the structure of the complex and the binding assays establish that the PAZ domain binds to the 3' end of RNA, requiring an essential 2-nt, single-stranded segment, and has a preference for the remaining RNA in the duplex over the single-stranded form.

The PAZ domain adopts a topology similar to the oligonucleotide/oligosaccharide-binding (OB) fold, despite different connections linking secondary structure elements^{13,14}. The 2-nt overhangs are anchored within a pocket unique to the PAZ domain, with the remainder of the RNA positioned along a surface formed by the $\beta 2$ - $\beta 3$ strands. This contrasts with OB fold RNA recognition, where the RNA targets the opposite face of the β -barrel¹⁸, corresponding to $\beta 6$ - $\beta 7$ strands of the PAZ domain.

The conservation among PAZ homologues, especially for the RNA-binding residues (Supplementary Fig. S3), predicts that all PAZ family members should similarly recognize siRNA ends, satisfying their exclusive distribution in the RNAi-related proteins Dicer and Argonaute. As an siRNA binder, Dicer-PAZ could hold the product siRNA at one end, after dsRNA cleavage, and transfer the other siRNA end to the PAZ domain of Argonaute in the RISC. Such a transfer would lead to a direct loading of siRNA into RISC, consistent with a previous observation that the activities of Dicer and RISC are coupled through transfer of siRNA⁷. Because some Dicer proteins, such as *Drosophila* Dicer-2, do not contain a PAZ domain, alternative pathways are likely to exist, that mediate transfer of siRNA from Dicer to RISC. *Drosophila* R2D2, which binds to Dicer-2, has been proposed to facilitate loading of siRNA into RISC¹⁹.

In view of the predominant interaction of PAZ with one strand of siRNA, and the associated residual binding to ssRNA, the PAZ

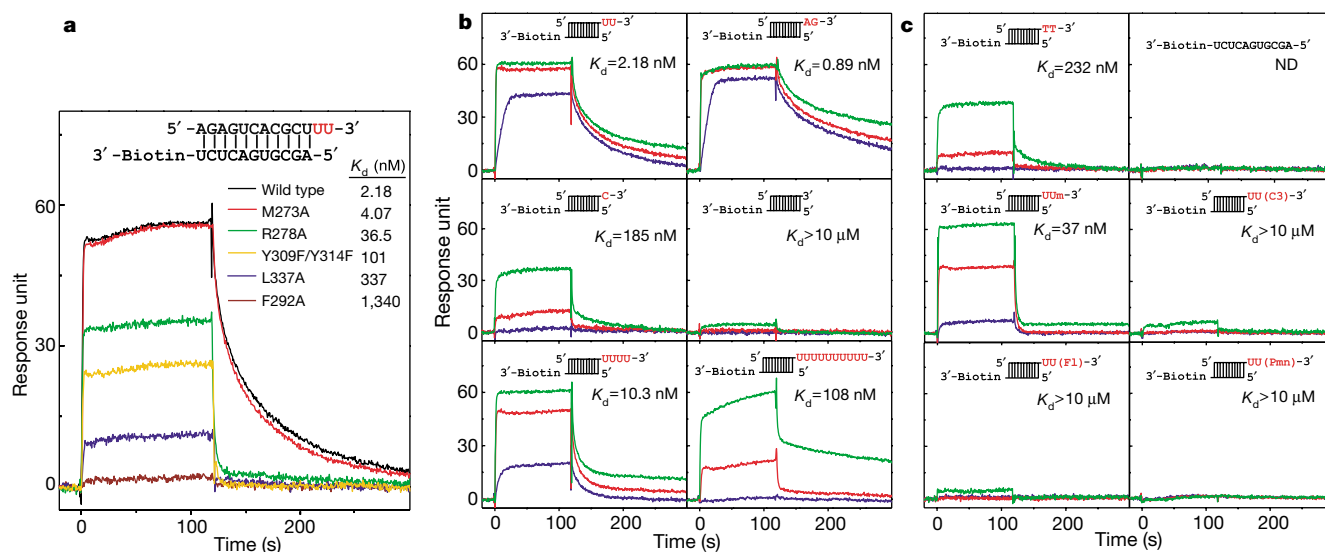


Figure 4 RNA-binding assays monitored using surface plasmon resonance.

a, Representative binding curves for wild-type and mutant PAZ domains (100 nM) and a UU 3'-overhang-containing RNA duplex. **b**, **c**, Representative binding curves for wild-type PAZ domain at concentrations of 10 nM (blue), 100 nM (red) and 1000 nM (green), and various RNA ligands. RNA ligands contain a common 11-bp duplex, shown explicitly

in **a** and schematically in **b** and **c**, which is biotin-labelled at the blunt end and contains either different 3' extensions (**b**) or 3'-end modifications (**c**) at the other end (highlighted in red). Um, 2'-*o*-methyl; C3, 3'-*o*-propanediol; F1, 3'-*o*-fluorescein; Pmn, 3'-puromycin. The association occurs during the first 120 s, and is followed by dissociation during the next 180 s. The deduced apparent K_d values are also listed.

domain might continue to secure the guide strand following siRNA unwinding, and present it for scanning and pairing with the target sequence in the activated RISC.

PAZ-siRNA recognition seems to provide the molecular basis for previous observations that RNAi efficiency was reduced when the 3' overhangs of synthetic siRNAs deviated from 2 nt in length^{7,15}, or when they were substituted by deoxyribonucleotide residues²⁰, or chemically modified at the 3'-OH position²¹. The PAZ domain represents the best candidate to sense these 3'-overhang modifications, which would result in reduced binding as shown experimentally in Fig. 4b, c, and consequently might impair the PAZ-mediated recruitment of siRNA into the RISC. By contrast, many other studies have reported that 3' modifications of siRNA^{11,22–24} and commonly used deoxynucleotide-substituted 3' overhangs¹⁵ do not interfere with RNAi, thereby questioning the importance of the PAZ-siRNA interaction in the RNAi pathway. One explanation is that PAZ may not be the only siRNA recognition domain mediating RNAi, with its contribution potentially compensated for by alternative high-affinity binding events during siRNA recruitment. Blunt-end siRNA can induce RNAi in some cases^{15,25}, probably by also by-passing PAZ-siRNA recognition. In addition, the nuclease stability of chemically modified siRNAs may partially compensate for impaired PAZ recognition. Further studies are needed to understand the full diversity of siRNA recognition events in the RNA silencing pathway. □

Methods

Protein and RNA preparation

The PAZ domain (residues 225–369) from human eIF2C1 was cloned into pET15b vector with a thrombin-cleavable N-terminal His tag. *Escherichia coli* strain BL21-Gold(DE3) harbouring the expression plasmid was induced by 1 mM isopropyl- β -D-thiogalactopyranoside and further grown overnight at 25 °C. After sonication-based breakage, the clarified cell lysate was mixed with 0.1% polyethylenimine to precipitate nucleic acid. The PAZ domain was purified by Ni-chelating and heparin Hitrap column, followed by His-tag removal with thrombin, and additional purification using Superdex 75 size-exclusion chromatography. Selenomethionine-labelled protein was cultured by inhibiting the methionine synthesis pathway. Site-directed mutations were made by a two-step PCR-based overlap extension method and confirmed by DNA sequencing. All proteins used in binding assays were purified without the removal of His tag and the use of

a final size-exclusion column. RNAs and their 3'-modified analogues were chemically synthesized, deprotected and purified by denaturing polyacrylamide gel electrophoresis, or purchased from Dharmacon.

Crystallization and structure determination

Crystals were grown by the hanging-drop vapour diffusion method at 20 °C. The initial drop consisted of 1 μ l of 10 mg ml⁻¹ PAZ and 5'-CGUGACUCU-3' 9-mer RNA, 100 mM KCl, 5 mM HEPES-KOH, pH 7.6, 10 mM dithiothreitol (DTT) and 1 μ l well solution of 20% PEG 1000, 50 mM HEPES-KOH, pH 7.6. Crystals were briefly cryoprotected in 15% PEG 1000 and 15% (w/v) sucrose (or 20% glycerol) before freezing in liquid nitrogen.

The diffraction data sets were collected at the Advanced Photon Source and processed using the HKL package²⁶. The crystals belong to the space group $P6_3$, with unit cell parameters of $a = b = 99.8$ Å, $c = 34.9$ Å, and contain 51% solvent. The asymmetric unit contains one-half of the complex (one PAZ monomer and one 9-mer ssRNA) with the crystallographic dyad coinciding with the dimeric axis in the complex.

The structure was solved by single anomalous diffraction (SAD) implemented in CNS²⁷ for the complex of selenomethionine-labelled PAZ and 9-mer RNA. After density modification by solvent flipping, the map calculated to 3.0 Å showed clear electron density for both protein and RNA in the complex (Fig. 3b). The model was built using the O program²⁸. Nevertheless, refinement was primarily undertaken against a better data set associated with a complex of PAZ and 9-mer RNA with the 2-nt, 3'-ribonucleotide overhang substituted by its deoxy counterpart. The refined PAZ-RNA/DNA complex structure, following conversion of the 2-nt 3' overhangs from deoxyribo to ribo residues, was used as the starting model for the refinement of the all-RNA structure. Both complex structures are essentially identical except for the conformation of the sugar puckers of the 2-nt 3' overhangs, which shift from 2'-endo pucker in deoxyribonucleotides to 3'-endo pucker in ribonucleotides. The conformation of sugar puckers was not constrained during refinement. The protein residues exhibit good stereochemistry with 85% located in the most favourable region, whereas the remaining fall in additional allowed regions of the Ramachandran plot. Detailed crystallographic statistics can be found in Supplementary Table 1. The symmetry-related guanines in the centre of the RNA duplex are tilted relative to each other, and fit well into the electron density. The final structures contain PAZ residues 225–295 and 302–349, 9 nucleotides and 15/16 water molecules. The figures were prepared with PyMOL (<http://pymol.sourceforge.net/>) and GRASP²⁹.

Biacore-based measurements of dissociation constants

Binding experiments³⁰ were performed on Biacore 2000 biosensor (Biacore) at 25 °C. The RNA constructs were designed to consist of an 11-bp duplex, with a biotin attached at one 3' end (blunt end) and different length 3' extensions at the opposite end. Our choice was based on the anticipation that PAZ would not bind to the blunt end and the short length of the duplex would minimize nonspecific complex formation. The 3' biotinylated 11-mer RNA (5'-AGCGUGACUCU-biotin-3') was first immobilized on a streptavidin-coated biosensor chip. The duplex was formed by injecting 500 nM of specific complementary strands in 1 M NaCl, followed by washing away unbound RNAs. The association was monitored over a 2-min period during which wild-type or mutant proteins, at concentrations ranging from 1 nM to 5 μ M and in a buffer of 0.01 M HEPES, pH 7.4,

0.15 M NaCl, 3 mM EDTA, 0.005% Surfactant P20 and 1 mM DTT, were flown over immobilized RNAs at a rate of $30 \mu\text{l min}^{-1}$, and dissociation monitored by flowing the blank buffer over a subsequent 3-min period. The residual bound protein was removed by injecting $20 \mu\text{l}$ of 2 M NaCl at $20 \mu\text{l min}^{-1}$. Experiments at each concentration of proteins were repeated at least twice. The apparent K_d values were derived from the global fit of the association and disassociation curves to a simple 1:1 Langmuir interaction model with a correction for mass transport effects using BIAevaluation 3.0 software.

Received 27 January; accepted 29 March 2004; doi:10.1038/nature02519.

- Denli, A. M. & Hannon, G. J. RNAi: an ever-growing puzzle. *Trends Biochem. Sci.* **28**, 196–201 (2003).
- Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
- Voinnet, O. RNA silencing as a plant immune system against viruses. *Trends Genet.* **17**, 449–459 (2001).
- Volpe, T. A. *et al.* Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833–1837 (2002).
- Hall, I. M. *et al.* Establishment and maintenance of a heterochromatin domain. *Science* **297**, 2232–2237 (2002).
- Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).
- Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**, 188–200 (2001).
- Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293–296 (2000).
- Hammond, S. M., Boettcher, S., Caudy, A. A., Kobayashi, R. & Hannon, G. J. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**, 1146–1150 (2001).
- Nykanen, A., Haley, B. & Zamore, P. D. ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107**, 309–321 (2001).
- Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R. & Tuschl, T. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563–574 (2002).
- Yan, K. S. *et al.* Structure and conserved RNA binding of the PAZ domain. *Nature* **426**, 468–474 (2003).
- Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain. *Nature* **426**, 465–469 (2003).
- Song, J. J. *et al.* The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nature Struct. Biol.* **10**, 1026–1032 (2003).
- Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W. & Tuschl, T. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J.* **20**, 6877–6888 (2001).
- Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
- Chiu, Y. L. & Rana, T. M. siRNA function in RNAi: a chemical modification analysis. *RNA* **9**, 1034–1048 (2003).
- Theobald, D. L., Mitton-Fry, R. M. & Wuttke, D. S. Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 115–133 (2003).
- Liu, Q. *et al.* R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway. *Science* **301**, 1921–1925 (2003).
- Hohjoh, H. RNA interference (RNAi) induction with various types of synthetic oligonucleotide duplexes in cultured human cells. *FEBS Lett.* **521**, 195–199 (2002).
- Harborth, J. *et al.* Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. *Antisense Nucleic Acid Drug Dev.* **13**, 83–105 (2003).
- Chiu, Y. L. & Rana, T. M. RNAi in human cells: basic structural and functional features of small interfering RNA. *Mol. Cell* **10**, 549–561 (2002).
- Holen, T., Amarzguioui, M., Wiiger, M. T., Babaie, E. & Prydz, H. Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Res.* **30**, 1757–1766 (2002).
- Amarzguioui, M., Holen, T., Babaie, E. & Prydz, H. Tolerance for mutations and chemical modifications in a siRNA. *Nucleic Acids Res.* **31**, 589–595 (2003).
- Czauderna, F. *et al.* Structural variations and stabilising modifications of synthetic siRNAs in mammalian cells. *Nucleic Acids Res.* **31**, 2705–2716 (2003).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).

- Jones, T. & Kjeldgaard, M. Electron-density map interpretation. *Methods Enzymol.* **227**, 174–208 (1997).
- Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).
- Katsamba, P. S., Park, S. & Laird-Offringa, I. A. Kinetic studies of RNA-protein interactions using surface plasmon resonance. *Methods* **26**, 95–104 (2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank K. Saigo for providing us with the eIF2C1 complementary DNA clone. This research was supported by the NIH. We thank Y. Cheng and personnel at the Advanced Photon Source (APS) beamlines 19BM and 14IDB for help in collecting the X-ray diffraction data. Use of the APS beamline was supported by the US Department of Energy, Basic Energy Sciences, Office of Science.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.J.P. (pateld@mskcc.org). Coordinates for the PAZ–siRNA complexes containing 2-nt ribo- and deoxyribonucleotide 3' overhangs have been deposited in the Protein Data Bank under accession codes 1SI3 and 1SI2, respectively.

retraction

Unaltered cosmic spherules in a 1.4-Gyr-old sandstone from Finland

Alexander Deutsch, Ansgar Greshake, Lauri J. Pesonen & Pekka Pihlaja

Nature **395**, 146–148 (1998).

We earlier reported finding unaltered cosmic spherules in samples from the Satakunta sandstone, Finland¹. At that time, we had found 18 spherules in 5 kg of material, and an additional 100 spherules were recovered from the 60–125 m heavy mineral separates of the samples. Rocks collected later from the same and nearby sites had no spherules in them². We have subsequently concluded that the cosmic spherules¹ were not part of the Satakunta sandstone samples. The total number of spherules we found during processing in the Muenster laboratory (about 120) is in itself a compelling argument for an exceptional contamination. □

- Deutsch, A., Greshake, A., Pesonen, L. & Pihlaja, P. Unaltered cosmic spherules in a 1.4-Gyr-old sandstone from Finland. *Nature* **395**, 146–148 (1998).
- Kettrup, D. *Mikrometeorite in terrestrischen Sedimenten—Eine systematische Analyse potentieller Wirtsgesteine unter besonderer Berücksichtigung des mesoproterozoischen Satakunta-Sandsteins (SW-Finnland)*. Inaug.-Diss., Univ. Münster (2002).

Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

US Head Office, New York

345 Park Avenue South, 10th Floor,

New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: naturejobs@natureny.com

US Sales Manager/ Corporations:

Peter Bless

Classified Sales Representatives:

Phone: +1 800 989 7718

Pennsylvania, New York,

Maryland/ Latin America/ NIH:

Michael McGillion

Midwest USA:

Wade Tucker

East USA/ Canada:

Janine Taormina

San Francisco Office

Classified Sales Representative:

Michaela Bjorkman

West USA/ West Corp. Canada

225 Bush Street, Suite 1453

San Francisco, CA 94104

Tel +1 415 781 3803

Fax +1 415 781 3805

e-mail: m.bjorkman@nature.sf.com

European Head Office, London

The Macmillan Building, 4 Crinan Street,

London N1 9XW, UK

Tel +44 (0) 20 7843 4961

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs Sales Director:

Nevin Bayoumi (4978)

Production Manager: Billie Franklin

To send materials use London

address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Stefan Hales

European Satellite Office

Germany/ Austria/ Italy/

The Netherlands/ Belgium:

Patrick Phelan, Odo Wulffen

e-mail: p.phelan@nature.com

o.wulffen@nature.com

Japan Head Office, Tokyo

MG Ichigaya Building (5F),

19-1 Harakatamachi,

Shinjuku-ku,

Tokyo 162-0841

Tel +81 3 3267 8751

Fax +81 3 3267 8746

Asia-Pacific Sales Director:

Rinoko Asami

e-mail: rasami@naturejpn.com

naturejobs

A delicate balance

When ten new member nations officially became part of the European Union (EU) earlier this month, it added a huge number of new recruits — and opportunities — to the region's scientific job market. But this expansion will require recruiters across the EU to strike a somewhat delicate balance. How, for example, can the wealthier EU nations in the west absorb thousands of scientists from the poorer new members without stretching their own resources and without draining the eastern countries of their intellectual capital?

The west should certainly prove to be a big draw for researchers in the east. The Baltic nations have a tradition of rigorous training in maths, physics and chemistry. These skills are in short supply in the west, so promising students could well find themselves wooed with promises of more funds and better infrastructure. Indeed, the west has regularly looked east for such skills, but EU membership changes the rules for visas and improves mobility.

This increased mobility carries with it the threat of an east-west brain drain. With some prudence and creativity, this could be avoided. One approach is to set up networks, such as ScanBalt, that emphasize collaborations between the east and west. Alternatively, efforts that encourage researchers within the eastern nations can help to persuade them to stay at home. For example, last summer, the medicinal-chemistry department at the University of Szeged in Hungary held a symposium that paired undergraduates with more senior scientists in a bid to give the youngsters more chances to publish.

Sadly, it is inevitable that many young scientists from the new member states will head west to seek their fortune. Hopefully, this will lead to a strengthening of infrastructure and investment in the new members that will eventually draw them, and others, east.

Paul Smaglik
Naturejobs editor



Contents

POSTDOCS AND STUDENTS

Making the switch between industry and academia p324

CAREER VIEW

Bricks & Mortar

Switzerland switches on supercomputers

Graduate Journal

A lab affair

Movers

Norbert Jousten

p326

WWW.NATUREJOBS.COM

Career centre

Information on the scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

ANNOUNCEMENTS

EVENTS

From industry to academia

A postdoc stint in industry can mean better pay and superior facilities. But will doing such a fellowship jeopardize your academic aspirations? Beth Martin investigates.

When Celia Schiffer was finishing her postdoc at South San Francisco biotech firm Genentech, she applied for both academic and industry positions. Along the way, she discovered the enormous difference between the two application processes. In the end she opted for academia; she is now an associate professor in the department of biochemistry and molecular pharmacology at the University of Massachusetts Medical School. In making the switch, she learned how to tailor her application for each sector — and prove that an industrial postdoc needn't be a hindrance on an academic career.

Most PhDs take a postdoc position in industry as an introduction to the private sector's research and development culture. This is an excellent deal for companies, who often hire postdocs as staff scientists, and gives the postdoc a head start in industry (see *Naturejobs* 5; 5 September 2002).

Others may do a postdoc in an industrial lab because they are drawn to a particular scientist or project, even if they do not feel committed to staying in industry. Some will then move back to university, where they can delve deeper into basic research and enjoy the autonomy of running their own labs.

The number of industry postdocs who go on to academia vary by industry, company and region. The picture in Europe is different from that in the United States, with fewer companies offering organized postdoc training programmes, instead opting to host individual Marie Curie fellows. Even with the Marie Curie programme, not many small biotech companies in Europe feel they have the resources to train fellows (see *Nature* 421, 296–297; 2003).

Big drug companies in Europe offer some postdoc positions, but there is even less mobility from there to academia than in the United States. In the United Kingdom, David Alker, discovery staffing group manager at Pfizer Global Research and Development in Sandwich, Kent, says that few young industry researchers return to universities. Pfizer has openings for four postdocs in its current discovery group, and he predicts that most will stay in industry, partly because they earn a lot more.

ACADEMIC ADVANTAGES

People who go back to academia want more ownership of their research and accept lower funding and support as a trade-off, says Alker. Overall, he believes that postdocs receive roughly the same amount of training in either sector.

Bart de Vos, who oversees about 70 research postdocs at Genentech, says that those who go on to seek university jobs may find it harder to understand the academic process after an absence of up to four

years. "If they don't stay in contact with their peers and advisers, they get out of touch with the academic culture," say de Vos. Anticipating that potential problem, Schiffer maintained relationships with her colleagues and advisers at the University of California, San Francisco, where she earned her PhD.

Another answer is networking, which is easier in some locations than others. For example, Cellzome is located very close to the European Molecular Biology Laboratory campus and its staff often go to the lab to attend seminars. Employees at companies in La Jolla, California, can walk freely into daily seminars at the Salk Institute, the Scripps Research Institute and the University of California, San Diego, among others. If the company doesn't stay so cosy with a university, it is up to the individual to stay in touch with former



Biotech firm Genentech gave Celia Schiffer valuable experience of industry.

GENENTECH

advisers and colleagues, and to contact academic researchers who did postdocs in industry to ask for advice.

But contacts alone won't help you make the switch. You also have to understand the underlying rhythm and culture of academia. Schiffer learnt that academic recruiting in North America operates mainly during a 'season', with most applications accepted from August to about February, and interviews occurring primarily between October and March, with the position typically starting the following autumn or winter. Job-seekers need to plan carefully, to ensure that their application materials are complete on time.

Understanding academia goes beyond scheduling, adds Paul Heyworth, senior scientist and chair of the postdoc programme at DNAX Research in Palo Alto, California. DNAX, the research division of Schering-Plough, supports 24 postdocs for three years of training. Industry postdocs are less experienced in writing grants and fellowship applications than their peers in academia, says Heyworth.

Schiffer realized that this was her weak point during the first year she submitted applications to universities.



Celia Schiffer: a two-year job hunt taught her to be self-critical while not taking rejection personally — she ended up with a choice of five posts.

Her research statement was originally based on the project she had been working on at Genentech. However, she had thought of working on HIV protease, so she contacted the leaders in the field — some of them former colleagues, some as cold calls — to help her work out where her research idea would fit.

FINDING THE RIGHT FIT

Now that Schiffer sits on her department's search committee and sees applications from the other side, she has a better understanding of how candidates can 'fit', whether they're from academia or industry. Her committee seeks a strong track record of publications and scientific achievement, but they also want to see a well-thought-out research plan, and evidence that the applicant understands its strengths and weaknesses.

Also, Schiffer adds, don't assume that coming from industry will count against you. Many industry postdocs have resources that their peers at university lack, such as ready access to equipment and reagents, which speeds the progress of their research.

Although a strong publishing record is valuable, search committees understand that industry postdocs could face more restrictions to publication. Jonathan Dantzig, who is on the search committee of the department of mechanical and industrial engineering at the University of Illinois, Urbana-Champaign, says that applicants from industry typically have fewer publications but more patents. The shorter list of publications won't count against you, he warns them, but your patents won't help as much as you hope. It is hard to measure any individual's contribution to a patent, he says, and very specialized knowledge may be needed to appreciate a patent's scientific merits.

Dantzig notes that candidates from industry bring valuable skills and experience. They have learned how

to conduct practical research as part of a team and how to manage projects. Their research often addresses a wider set of problems and they have interacted with scientists from other disciplines. Along with their broader perspective, they bring industry contacts that are useful in establishing collaborations, especially for engineering and translational biomedical research.

"What I learned in industry made me more successful in the university setting," says Dantzig, who spent four years working for an engineering company before returning to university.

Schiffer agrees. She says her experience at Genentech opened her eyes to new avenues of research and taught her how to run an experimental lab.

PERSISTENCE PAYS

Schiffer's top tips are to be persistent, not to take things personally, and to be self-critical. She spent two years making applications before finding her job at the University of Massachusetts. The first year, she says, "I submitted to 40 places, got four interviews, and no offers." So she spent the next year honing her research statement and learning how to sell herself.

The second year, she submitted 120 applications, got 11 interviews, and five offers. Once she was hired, her first five grant applications were rejected. After six years, everything is running smoothly, with a lab of ten people and three NIH grants. This process is normal, she says, although she concedes that the sheer volume of her applications might be exceptional. "The rejection letters," she says, "can make nice wallpaper."

In the end, it comes down to pursuing your goal and not getting discouraged. It might take some extra effort, but overall, doing your postdoc in industry puts you at no disadvantage in landing a university job. ■

Beth Martin is a freelance writer based in San Francisco, California.

GRADUATE JOURNAL

A lab affair

I did not expect to find myself in this kind of relationship. We met when the new arrival moved into a lab upstairs. It was difficult for us to communicate at first, but after spending a week together, I returned to my lab with a deeper understanding of the possibilities and pitfalls ahead of us.

My new labmate and close companion, a powder diffractometer, would take me to new heights in identifying and characterizing powdered crystalline material. On making this instrument's acquaintance, my last mechanical flame, an older model from the geology department, faded into distant memory.

Could it be a sense of possessiveness that rises in me when another student asks to use it? How is it that I feel guilty when I forget to run a weekly calibration? Why would anger flash when others fail to clean up when their run is over?

I can't help but anthropomorphize this bucket of bolts. Too many joys, sorrows and blank stares have been shared. Too many troubles were overcome and diaries written about our relationship not to admit that I am close to, and will certainly miss, my diffractometer when I go. Without getting to know this analytical tool, generating the new insights into my material on the Ångström scale would not have been possible. ■

Sidney Omelon is a PhD student in bone biomaterials at the Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Canada.

BRICKS & MORTAR

Switzerland switches on supercomputers

Last month, a supercomputing centre devoted solely to bioinformatics came online in Switzerland. Based at the Swiss Institute of Bioinformatics in Lausanne, Vital-IT can process 450 billion calculations per second and should provide a welcome boost to the country's proteomics capabilities.

As well as raising Switzerland's profile in the world of bioinformatics, Vital-IT will create several research jobs and will help to foster innovation. The centre is a collaboration between the bioinformatics institute, the computing firms Hewlett Packard and Intel, and five other academic institutions: the Ludwig Institute for Cancer Research, the Swiss Federal Institute of Technology (EPFL) in Lausanne, and the universities of

Lausanne, Geneva and Basel. All of the partners will benefit from having access to the new facility. The EPFL, for example, is setting up two research groups in bioinformatics that will be committed to the centre.

The Swiss government this year doubled the budget of the Swiss Institute of Bioinformatics to SFr6 million (US\$4.6 million) a year starting in 2004; the institute will have received a total of SFr27 million by 2007. About half of that money will go towards paying for protein annotation, the labour-intensive process of determining each protein's function and classifying them by family. The budget hike has allowed Vital-IT to hire five staff this year and may provide for more positions next year, says Victor Jongeneel, the centre's director.

Hewlett Packard and Intel got involved in Vital-IT primarily as a means of

testing new technologies and as a way to get into the potentially lucrative bioinformatics market. Their efforts seem to be paying off: the centre is already attracting talented computer scientists to tweak old bioinformatics software to run on the new hardware (see *Nature Biotechnol.* **22**, 492–493; 2004). The two companies, which are funding one position each at the centre, are also building their own in-house bioinformatics teams, says Jongeneel, who adds that all academic partners in the project intend to grow their research capacity over the next few years.

But, perhaps most importantly, the launch of the centre should attract more collaborators who are interested in applying new computational techniques to protein analysis. ■

Paul Smaglik is editor of Naturejobs.
♦ www.isb-sib.ch/projects/vitalit.htm

MOVERS

Norbert Jousten, executive director, International Science and Technology Center, Moscow



With his affinity for eastern Europe, his diplomatic abilities and his solid background in nuclear technology, Norbert Jousten feels prepared for his new post as executive director of the International Science and Technology Center (ISTC) in Moscow.

Jousten is used to change. In December 2000, he was a diplomat for the European Commission in the

Ukrainian city of Kiev when the government there submitted to the will of the international community and closed the last nuclear reactor at Chernobyl. "It was a fascinating time," he remembers. "You could feel the tremendous changes going on and the great desire of the people to come closer to Europe."

Now the 57-year-old Belgian is preparing to help bring about further changes in the eastern bloc. The ISTC helps to make peaceful use of the legacy of the former Soviet Union's nuclear programme. The end of the cold war left many highly trained weapons experts without a job. Since 1994, the ISTC has been scouring their home countries in an attempt to find them civilian research jobs.

Jousten, who studied nuclear physics in Liège, Belgium, before beginning his career in nuclear technology and, later, diplomacy, is optimistic about the prospects of success in his newest mission. So far, he

notes, fears that former Soviet weapons scientists could be tempted to offer classified knowledge to 'undesirable' paymasters have proved unfounded — at least, in most cases. Jousten hopes that, as Russia's economic situation stabilizes, the ISTC's attempts to redirect military know-how into socially beneficial activities will continue to prevent scientists from defecting. "There are many newly emerging opportunities for military researchers," he says.

Currently, Jousten is living without his wife and family in a Moscow hotel, where he is trying to improve his Russian. "Not an easy task," he admits. "But Moscow is such an incredibly fast-developing city, I really like working here." By the end of October, when the ISTC will celebrate its tenth anniversary with a big conference, Jousten hopes that the search for a flat will be over, his Russian will be good enough to order dinner, and his family will be reunited. ■

CV

1996–2004: Head of European Commission delegation to Ukraine, Belarus and Moldova in Kiev, Ukraine

1992–96: Head of Regional Cooperation and Nuclear Safety unit at the European Commission

1988–92: European Union representative at the International Atomic Energy Agency, Vienna, Austria

1980–88: Safety inspector, EURATOM, Brussels, Belgium

1975–79: Nuclear automation engineer, Westinghouse, Belgium

1970–75: Nuclear power plant engineer, Siemens, Erlangen, Germany